

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).

Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

THIS PAGE LEFT BLANK

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

THIS PAGE LEFT BLANK

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 21:23:21 ; Search time 92.5884 Seconds
(without alignments)
4645.181 Million cell updates/sec

Title: US-10-728-323-1
Perfect score: 3586
Sequence: 1 aataatcatatattatc.....cggttgctggtgtttctcc 2032

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_n2p.model -DEV=xlp
-Q=/abs/ABSWEB_spool/US10728323/runat_15052006_172132_22377/app_query.fasta_1
-DB=UniProt -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss02p
-USER=US10728323 @CGN_1_1_694 @runat_15052006_172132_22377 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOK=100 -LONLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3286	91.6	626	1 ALL12_ARAHY	P43238 arachis hyp
2	3286	91.6	626	2 Q547W5_ARAHY	Q547W5 arachis hyp
3	3041	84.8	614	1 ALL11_ARAHY	P43237 arachis hyp
4	2887.5	80.5	580	2 Q6PSU3_ARAHY	Q6PSU3 arachis hyp
5	2172.5	60.6	428	2 Q6PSU4_ARAHY	Q6PSU4 arachis hyp
6	1495.5	41.7	299	2 Q6PSU5_ARAHY	Q6PSU5 arachis hyp
7	1435	40.0	303	2 Q6PSU6_ARAHY	Q6PSU6 arachis hyp
8	1302	36.3	621	2 Q7XXT2_SOYBN	Q7XXT2 glycine max
9	1297	36.2	621	2 Q4LER6_SOYBN	Q4LER6 glycine max
10	1290	36.0	533	2 Q6EBC1_LUPAL	Q6EBC1 lupinus alb
11	1286	35.9	604	2 Q4LER5_SOYBN	Q4LER5 glycine max
12	1284	35.8	621	2 Q94BY0_SOYBN	Q94BY0 glycine max
13	1282	35.8	605	2 Q94IX2_SOYBN	Q94IX2 glycine max
14	1281	35.7	605	1 GLCA_SOYBN	P13916 glycine max
15	1267	35.3	571	1 CVCA_PEA	P13915 pisum sativ
16	1265.5	35.3	623	2 Q948X9_SOYBN	Q948X9 glycine max

ALIGNMENTS

RESULT 1

ALL12_ARAHY

ID ALL12_ARAHY STANDARD; PRT; 626 AA.

AC P43238;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE Allergen Ara h 1, clone P41B precursor (Ara h 1).

OS Arachis hypogaea (Peanut).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;

OC Arachis.

OX NCBI_taxID=3818;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=cv. Florunner;

RX MEDLINE=96013631; PubMed=7560062;

RA Burks A.W., Cockrell G., Stanley J.S., Helm R.M., Bannon G.A.;

RT "Recombinant peanut allergen Ara h I expression and IGE binding in

patients with peanut hypersensitivity.";

RL J. Clin. Invest. 96:1715-1721(1995).

RN [2]

RP CARBOHYDRATE-LINKAGE SITE ASN-521.

RX MEDLINE=20455243; PubMed=10998264; DOI=10.1006/abio.2000.4737;

RA Kolarich D., Altmann F.;

RT "N-glycan analysis by matrix-assisted laser desorption/ionization mass

spectrometry of electrophoretically separated nonmammalian proteins:

application to peanut allergen Ara h 1 and olive pollen allergen Ole e

1.";

RL Anal. Biochem. 285:64-75(2000).

CC -!- ALLERGEN: Causes an allergic reaction in human.

CC -!- SIMILARITY: Belongs to the 7S seed storage protein family.

CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----

DR EMBL; L34402; AAB00861.1; -; mRNA.
DR HSP; P25974; IIPJ.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR Pfam; PF00190; Cupin.1; 2_
KW Allergen; Glycoprotein; Signal.
FT SIGNAL 1 25 Potential.
FT CHAIN 26 626 Allergen Ara h 1, clone P41B.
FT CARBOHYD 521 521 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 626 AA; 71345 MW; 1A6BBE4149D0E3 CRC64;

Alignment Scores:

Pred. No.: 1,66e-208 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.6% Indels: 0
DB: 1 Gaps: 0

US-10-728-323-1 (1-2032) x ALL12_ARAHY (1-626)

QY 50 ATGAGAGGAGGGTTTCTCCACTGATCTGTGTCTAGGGATCCTTGTCTCGGCTTCAGTT 109
DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyileLeuValLeuA1aSerVal 20
QY 110 TCTGCAACGATGCCAAGTATCATCATTACAGAGAAACACAGAGAACCCCTGGCCAG 169
DB 21 SerAlaThrHisAlaIysSerSerProTyrGlnLysLysThrGluAenProCysAlaGln 40
QY 170 AGTGCCCTCAGAGTGTCTAACAGGACCGATCTTGAAGCAAAAGGCATCGAGTCT 229
DB 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGTGTCAACCAAGTCTCGATGATCTCTGTGTCTATGATCTCTCAGGACACACTGGC 289
DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACACCAACCAAGTCTCCCTCCAGGGAGCGGACAGCTGGCCGCCAACCCGGAGACTAC 349
DB 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATCACCGCTCAACCCGAGAGAGAGAGGAGGCGCGATGGGACACAGCTGGACCG 409
DB 101 AspAspAspArgArgGlnProArgGluGluGlyGlyArgTyrProAlaGlyPro 120
QY 410 AGGAGCGGTGAAGAGAAGAACTGGAGACAACCAAGAGAAGATTGGAGCGGACCAAGT 469
DB 121 ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgArgProSer 140
QY 470 CATCAGAGCCACGGAAATAAGCCCGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGlyGluGlnGluTyrGlyThr 160
QY 530 CAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCTTCTACTTCCCGTCAAGG 589
DB 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCCACCCCTACGGGAACCAAAACGGTAGGATCCGGGTCTCGAGAGGTTTGAC 649
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgValLeuGlnArgPheAsp 200
QY 650 CAAGGTCAAGCGATTTCAAGATCTCCAGATCACCAGTATGTGGAGATCGAGGCCAAA 709
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTGTCTTCCCAAGCAGCTGATCTGATCAATCACTTGTATCCAGCAA 769
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTAGCGTAAGAAATGGCAATTAACAGAAAGACTTTAATCTTTGACGAG 829
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260

QY 830 GGCCATGCACACTCAGAAATCCATCCGCTTTCATTTCTACATCTTGAACCCCATGACAAC 889
DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACTCAGAGTAGTAAATCTCATGCCCGTTAAACACACCCCGCCAGTTTGGAGAT 949
DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGGAGAGCGAGACCAATCATCTCTACTTGCAGGCTTTCAGCAGGAATACG 1009
DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGCGCCCTTCAATCGGAATTCATAGATACGAGGCTGTGTGTAGAGAGAAT 1069
DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGAGGTGAGCAAGAGAGAGAGCGAGCGATGAGTACTCGGAGTACTGAGAAC 1129
DB 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGGATGATAGTCAAAAGTGTCAAAGGACACGTTGAAGAACTTACTAAGCAGCT 1189
DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCGTCTCAAAGAAAGCTCCGAGAGAGAGAGATATCAACCAACCAATCAACTTG 1249
DB 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGAGGCGAGCCCGATCTTCTTAAACACTTTTGGGAAGTATTTTGAGGTGAAGCCAG 1309
DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAACCCCGACTTCAGGACCTGGACATGATCTCTACCTGTGTAGAGATCAAGAA 1369
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GAGACTTGTATGCTCCACACTTCAACTCAAGGCGATGTTATCTGCTGCTCAACAAA 1429
DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAACCTGGAAACCTTGAACTCGTGGCTGTAAAGAAAAAGAGCAACACAGAGGGACGCGG 1489
DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnArgArg 480
QY 1490 GAAGAAGAGAGACGAAGACGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1549
DB 481 GluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTGAAGGAAGCGGATGTGTTTCATCATGCCAGCTATCCAGTAGCCATC 1609
DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGAACCTCCATCTGTGGCTTCGGTATCAACGCTGAAAAACAACACAGA 1669
DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTCAGGTGATAAGGCAATGTGATAGACCATAGACAGAGAGAGAGAGAGAT 1729
DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGTGGTGAACAGTGTGAGAGCTCATCAAAAACAGAGAGGAATCT 1789
DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTGTCTCTCTCAATCTCAATCTCAATCTCCGCTGTCTCTGAGAAAGAG 1849
DB 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerProGluLysGlu 600
QY 1850 TCTCTCTGAGAAAGAGATCAAGAGGAGGAGAAACCAAGAGGAGGAGGCTCCACTCTTCA 1909
DB 601 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTTAAC 1927

```
Db      621 11eLeuLysAlaPheAsn 626
RESULT 2
Q547W5_ ARAHY
ID   Q547W5_ARAHY PRELIMINARY;      PRT;   626 AA.
AC   Q547W5;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Major allergen Ara h1.
OS   Arachis hypogaea (Peanut).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC   rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC   Arachis.
OX   NCBI_TaxID=3818;
RN   [1];
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22960226; PubMed=14597159; DOI=10.1016/j.molimm.2003.09.002;
RA   Viquez O.M., Konan K.N., Dodo H.W.;
RT   "Structure and organization of the genomic clone of a major peanut
RT   allergen gene, Ara h 1";
RL   Mol. Immunol. 40:565-571(2003).
DR   EMBL; AF432231; AAL27476.1; -; Genomic DNA.
SQ   SEQUENCE 626 AA; 71345 MW; 1A6BBB41490D0E3 CRC64;

Alignment Scores:
Pred. No.:      1,66e-208      Length:      626
Score:          3286.00        Matches:      626
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%    Mismatches:  0
Query Match:     91.6%          Indels:       0
DB:              2             Gaps:         0

US-10-728-323-1 (1-2032) x Q547W5_ARAHY (1-626)
QY      50  ATGAGAGGAGGTTTCTCCACTGATGCTGTCTAGGATCCTGTCTCGTTCAGTT 109
Db      1  MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyLeuValLeuAlaSerVal 20
QY      110 TCTGCACCGATCCCAAGTCATCACCCTTACGAGAGAAACAGAGAACCCCTCGGCCGAG 169
Db      21  SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY      170 AGTGCTCCACAGATTGTCAACAGGAACCGGATGACTTCAAGCAAAAGGCATGCCAGTCT 229
Db      41  ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60
QY      230 CGTGCACCAAGCTCAGATGATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGGC 289
Db      61  ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY      290 ACCACCAACACAGTTTCCCTCCAGGGGAGGAGACACGTGGCGCGCAACCCGAGACTAC 349
Db      81  ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY      350 GATGATACCGCGTCAACCCCGAGAGAGAGGAGGCGGATGGGACCACTGGACCG 409
Db      101  AspAspAspArgGlnProArgGlnProArgGlnGluGlyArgTyrGlyProAlaGlyPro 120
QY      410 AGGGAGCGTGAAGAGAAGAAGACTGGAGACCAACCAAGAGAAGATTGGAGGCGCAAGT 469
Db      121  ArgGluArgGluArgGluGluAspTyrArgGlnProArgGluAspTyrArgProSer 140
QY      470 CATACAGCCGACGGAATAAGGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
Db      141  HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGlnTyrGlyThr 160
QY      530 CCAGGTAGCCATCTGAGGAGAGAGACATCTCGGACCAACCCCTTCTACTTCCCGTCAGG 589
Db      161  ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY      590 CGGTTTAGCACCCCGTACGGGAACCAAAACGGTAGATCCGGTCTCTGCAGAGGTTTGAC 649
Db      181  ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY      650 CAAAGGTCAAGGACGTTTCCAGAACTCCAGAACTACCGCTATTGTGCAGATCCAGGCCAAA 709
Db      201  GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY      710 CTTAACTACTTGTCTTCTCCCAAGCAGCTGATGTGTGATAACATCCTTGTATTCCAGCAA 769
Db      221  ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY      770 GGGCAAGCCACCGTGACCGTAGCAATAATGCAATAACAGAAAGAGCTTTAATCTTGACGAG 829
Db      241  GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY      830 GGCCATGCACTCAGAACTCCCATCCCGTTTCATTTCTTACATCTTGAACCGCCATGCAAC 889
Db      261  GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY      890 CAGAACCTCAGAGTAGTAAATCTCCATGCCGTTTAAACACACCCGCGCAGTTTGAGGAT 949
Db      281  GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY      950 TTCTTCCCGGAGCAGCGAGACCAATCATCTACTTTCAGCGGCTTCAGCAGGAATACG 1009
Db      301  PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY      1010 TTGGAGGCGGCTTCAATCCGGAATTCATGATGATACGGAGGCTGCTGTAGNAGAT 1069
Db      321  LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn 340
QY      1070 GCAGGAGGTGAGCAAGAGGAGAGGCGAGCGGATGAGTACTCGGAGTAGTGAGAAC 1129
Db      341  AlaGlyGlyGluGlnGluArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn 360
QY      1130 AATGAAGAGTGATAGTCAAGAGTCAAGGAGCAGCTTGAAGAACTTACTTAAGCAGCT 1189
Db      361  AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY      1190 AATCCGCTCTCAAGAAAGGCTCCCAAGAGGAGAGATATCACCACCCCAATCACTTG 1249
Db      381  LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY      1250 AGAGAAGCGAGCGCCGATCTTCTTAACTTTGGGAAGTATTATTGAGGTGAAGCGAGAC 1309
Db      401  ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY      1310 AAGAAAGACCCCGCTTCCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369
Db      421  LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY      1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGGTTATCGTCTGCTCAACAAA 1429
Db      441  GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY      1430 GSAACTGGAAACTCTTGAACCTCGTGGCTGTAAAGAAAGCAACAACAGAGGGGACGGCG 1489
Db      461  GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnArgGlyArgArg 480
QY      1490 GAAGAAGAGGAGGACGAAGCAAGAGAGGAGGAAGTAAACAGAGAGGTGCGTAGGTAC 1549
Db      481  GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY      1550 ACAGCGAGGTGAAGGAAGGCGATGTGTTTCATGTCGACGAGCTCATCCAGTAGCCATC 1609
Db      501  ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY      1610 AACGCTTCTCCGAACTCATCTGCTTGGCTTTCGCTATCAACCGCTGAAACCAACACACAG 1669
Db      521  AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY      1670 ATCTTCTTGCAGGTGATAAGGACAATGTGTAGACCAAGATAGAGAAAGCAAGCAGGAT 1729
```

```
Db      541  llePheLeuAlaGlyAspLysAspIleAspGlnIleGluLysGlnAlaLysAsp 560
QY      1730  TTAGCATTCCTCCGCGGTGAACAGTGTGAGAGCTCATCAAAACAGAGGAATCT 1789
Db      561  LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY      1790  CACTTTGTGAGTGTCTCTCATCTCAATCTCAATCTCCGTCGTCTCTCTGAGAAGAG 1849
Db      581  HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY      1850  TCTCCTGAGAAAGAGATCAAGAGGAGGAGAAACCAAGGAGGAGGAGGAGGAGGAGG 1909
Db      601  SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyGlyGlyProLeuLeuSer 620
QY      1910  ATTTTGAAGCTTTTAAAC 1927
Db      621  IleLeuLysAlaPheAsn 626

RESULT 3
ALL11 ARAHY STANDARD; PRT; 614 AA.
AC P43237;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Allergen Ara h 1, clone p17 precursor (Ara h 1).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eutrosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OC NCBI_TaxID=3818;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=cv. Florunner;
RX MEDLINE=96013631; PubMed=7560062;
RA Burks A.W., Cockrell G., Stanley J.S., Helm R.M., Bannon G.A.;
RT "Recombinant peanut allergen Ara h 1 expression and IgE binding in
RT patients with peanut hypersensitivity.";
RN J. Clin. Invest. 96:1715-1721(1995).
RN [2]
RP CARBOHYDRATE-LINKAGE SITE ASN-516.
RX MEDLINE=2045243; PubMed=10998264; DOI=10.1006/abio.2000.4737;
RA Kolarich D., Altman F.;
RT "N-glycan analysis by matrix-assisted laser desorption/ionization mass
RT spectrometry of electrophoretically separated nonmammalian proteins:
RT application to peanut allergen Ara h 1 and olive pollen allergen Ole e
RT 1.";
RL Anal. Biochem. 285:64-75(2000).
CC -1- ALLERGEN: Causes an allergic reaction in human.
CC -1- SIMILARITY: Belongs to the 7S seed storage protein family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; L38853; AAA60336.1; -; mRNA.
DR HSSP; P25974; 1IPJ.
DR IntAct; P43237; -.
DR GlycoSuiteDB; P43237; -.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR Pfam; PF00190; Cupin.1; 2.
KW Allergen, Glycoprotein; Signal.
FT SIGNAL 1 25 Potential.
FT CHAIN 26 614 Allergen Ara h 1, clone P17.
FT CARBOHYD 516 516 N-linked (GlcNAc...)/FTid=CAR_000218.
FT SEQUENCE 614 AA; 70283 MW; 1DDACF2I7ECSF31 CRC64;
```

```
Alignment Scores:
Pred. No.: 2,57e-192 Length: 614
Score: 3041.00 Matches: 595
Percent Similarity: 96.0% Conservative: 9
Best Local Similarity: 94.6% Mismatches: 7
Query Match: 84.8% Indels: 18
DB: 1 Gaps: 7

US-10-728-323-1 (1-2032) x ALL11_ARAHY (1-614)
QY 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTGTGTAGGATCCTTGTCTCGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGCATGCCAAGTCATCCTTACCAGAGAGAAACAGAGAACCCCTGCGCCAG 169
Db 21 SerAlaThrGlnAlaLys---SerProTyr---ArgLysThrGluAsnProCysAlaGln 38
QY 170 AGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAACAAAGGCGATCGGAGTCT 229
Db 39 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 58
QY 230 CGCTGCACCAAGCTCGAGTATGATCCTCGTTGTCTATGATCTCGAGGACACACTGCG 289
Db 59 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAsp-----ThrGly 74
QY 290 ACACCAACCAACGTTCCCTCCAGGGAGCGGACACGTCGGCGGCCCAACCCGAGACTAC 349
Db 75 AlaThrAsnGlnArgHisProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 94
QY 350 GATGATGACCGCGCTCAACCCCGAAGAGAGAGGAGGCGCGATGGGACCCAGCTGACCG 409
Db 95 AspAspAspArgArgGlnProArgGluGluGlyArgTyrGlyProAlaGluPro 114
QY 410 AGCGAGCGTGAAGAGAGAAAGACACTCGAGACAAACCAAGAGAGATTGGAGGGCCAAAGT 469
Db 115 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 134
QY 470 CATCAGCAGCCAGGAAATAAGGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
Db 135 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 154
QY 530 CAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACCAACCTTCTTACTTCCCGTCAAGG 589
Db 155 ProGlySerGluValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 174
QY 590 CGGTTTAGCACCCTGCTACCGGAAACCAAAACGGTAGGATCCGGTCTCTCGAGAGTTTGAC 649
Db 175 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 194
QY 650 CAAAGTCAAGGCAGTTTCAGAAATCTCCAGAAATCCAGTATTGTGCAGATCCGAGGCCAA 709
Db 195 GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg 214
QY 710 CCTAACACTCTTGTCTTCTCCCAAGCACGCTGATGCTGTGATAACATCTTGTATCCAGCAA 769
Db 215 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGln 234
QY 770 GGGCAAGCCACCGTACCGTAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 829
Db 235 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 254
QY 830 GGCATGCACTCAGAAATCCATCCGTTTCATTTTCATTTTCATTTTCATTTTCATTTTC 889
Db 255 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 274
QY 890 CAGAACCTCAGAGTAGCTAAATCTCATGCCCGTTTAAACACACCCCGCCAGTTTGGAGAT 949
Db 275 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 294
QY 950 TTCTTCGCGGAGCAGCGAGACCAATCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1009
Db 295 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 314
```


QY	650	CAAAGTCAAGGCGAGTTTCAGAAATCTCCAGAAATCACCGCTATTGTGCGAGATCGAGGCCAAA	709
Db	195	GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg	214
QY	710	CCTAACACACTCTGTTCTCCCAAGCAGCGTGATGCTGATAACATCTCTGTTATCCAGCAA	769
Db	215	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	234
QY	770	GGCAAGCCACCGTAGCCGTAGCAAAATGGCAATTAACAGAAAGAGCTTTAATCTTGACGAG	829
Db	235	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	254
QY	830	GGCCATGCACTCAGAAATCCCATCCGTTTCATTTCTCTACATCTTGAAACCGCCCATGACAAC	889
Db	255	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	274
QY	890	CAGAACTCAGACTAGCTAAATCTCATGCCCGTTAAACACACCCGCCGAGTTTGAGGAT	949
Db	275	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	294
QY	950	TTCTTCCCGCAGCAGCCGAGCAATCATCTACTTGCAGGCTTCAGCAGGAATACG	1009
Db	295	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	314
QY	1010	TTGAGGCGCCCTTCAATGCGGAATTCATGAGATACGGAGGTGCTGTTAGAGAAGAT	1069
Db	315	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn	334
QY	1070	GCAGGAGGTGACCAAGGAGGAGCGCAGCGATGAGTACTCGGAGTACTGAGAAC	1129
Db	335	AlaGlyGlyGlnGluGluArgGlyGlnArgArgSerThrArgSerSerAsp---	353
QY	1130	AATCAAGGAGTGATAGTCAAAAGTGTCAAAGGACACGTTGAAGAACTTACTAAGCAGCT	1189
Db	354	AsnGluGlyValIleValLysValSerLysGluHisValGlnGluLeuThrLysHisAla	373
QY	1190	AAATCCGTCTCAAGAAAGCTCCGAAGAGAGGGAGATATCAACCAACCAATCAACTTG	1249
Db	374	LysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu	392
QY	1250	AGAGAAGGCGAGCCGCATCTTTCAACACTTTGGGAAGTTATTGAGGTGAAGCCAGAC	1309
Db	393	ArgAspGlyGluProAspLeuSerAsnAsnPheLysArgLeuPheGluValLysProAsp	412
QY	1310	AAGAAGAACCCCGAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA	1369
Db	413	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu	432
QY	1370	GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCGTCGTCGTCACAAA	1429
Db	433	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	452
QY	1430	GGAACCTGGAACCTTGAACTCGTGCTTAAGAAAAGAGCAACAACAGAGGGGACGCGG	1489
Db	453	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnArgGlyArgArg	472
QY	1490	GAA-----GAAGAGGAGCGAAGACGAAGAAGAGGAGGGAAGTAACAGAGAGGTG	1540
Db	473	GluGlnGlnrPpGluGluGluGluGluAspGluGluGluGlySerAsnArgGluVal	492
QY	1541	CGTAGGTACACAGGAGGTGAAGGAAGGCGATGTTTCATCATGCCAGAGCTCATCCA	1600
Db	493	ArgArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisPro	512
QY	1601	GTAGCCATCAACGCTTCTCCGAACCTCAATCTGCTGGCTTCGGTATCAACGCTGAAAC	1660
Db	513	ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn	532
QY	1661	AACCACAGAAATCTTCTGCGAGTGTAAAGCAATGTGTATAGACCAGATAGAGAAGCAA	1720
Db	533	AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln	552
QY	1721	GCAGAGGATTTAGCATTCCTCGGTCGGCTGAACAGTTGAGAAGCTCAAAAAACCG	1780
Db	553	AlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGln	572
QY	1781	AAGGAATCTCAGCTTTTGTGAGTGCT	1804
Db	573	ArgGluSerHisPheValSerAla	580
RESULT 5			
ID	Q6PSU4	ARAHY	
AC	Q6PSU4	ARAHY PRELIMINARY;	PRT; 428 AA.
DT	05-JUL-2004	(TrEMBLrel. 27, Created)	
DT	05-JUL-2004	(TrEMBLrel. 27, Last sequence update)	
DE	05-JUL-2004	(TrEMBLrel. 27, Last annotation update)	
OS	Arachis hypogaea (peanut).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eurosid1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;		
OC	Arachis.		
OX	NCBI_TaxID=3818;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RA	Wang L., Liao B., Li H., Yan Y., Lin X., Huang S.;		
RL	Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY581851; AAR00596.1; -; mRNA.		
DR	GO; GO:0045735; F:nutrient reservoir activity; IEA.		
DR	InterPro; IPR006045; Cupin.		
DR	InterPro; IPR007113; Cupin_region.		
DR	Pfam; PF00190; Cupin_1; 2.		
FT	NON_TER		
SQ	SEQUENCE	428 AA; 48094 MW; F1A32DAEC33AD06F CRC64;	
Alignment Scores:			
Pred. No.:	5.85e-135	Length:	428
Score:	2172.50	Matches:	427
Percent Similarity:	99.8%	Conservative:	0
Best Local Similarity:	99.8%	Mismatches:	0
Query Match:	60.6%	Indels:	1
DB:	2	Gaps:	1
US-10-728-323-1 (1-2032) x Q6PSU4_ARAHY (1-428)			
QY	644	TTTGACCAAGTCAAGGAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCGAGTCGAG	703
Db	2	PheAspGlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGlu	21
QY	704	GCCAAACCTTAACACTCTGTTCTTCCCAAGCAGCGTGATGCTGATAACATCCTTGTATC	763
Db	22	AlaLysProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIle	41
QY	764	CAGCAAGGCGCAAGCCACCGTAGCAATGGCAATTAACAGAAAGAGCTTTAATCTTT	823
Db	42	GlnGlnGlyGlnAlaThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeu	61
QY	824	GAGAGGGCCATGCATCCTCAGAAATCCCATCCGTTTTCATTCCTACATCTTGACCCCAT	883
Db	62	AspGluGlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHis	81
QY	884	GACAACGAGAACTCAGAGTAGCTAAATCTCCATGCCGTTAAACACACCCGCCAGTTT	943
Db	82	AspAsnGlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPhe	101
QY	944	GAGGATTTCTTCCCGCGCAGCAGCCGAGACCAATCATCTTCTTCTGAGGCGTTTCAGCAGG	1003
Db	102	GluAspPhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArg	121
QY	1004	AATACGTTGAGGCGCCCTTCATTCGGAATTCATAGATAGACGAGGCTGCTGTAGAA	1063
Db	122	AsnThrLeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuGlu	141
QY	1064	GAGAATGCAGGAGGTGAGCAAGAGGAGAGAGGCGCAGATGAGTACTCGGAGTAGT	1123

Db 142 GluAsnAlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTyrSerThrArgSerSer 161
QY 1124 GAGAAATGAGAGAGTGTACTGCAAGCTGCAAGAGGACCGTTCAGAACTTACTAAG 1183
Db 162 GluAsnAsnGluGlyValIleValIleValIleValIleValIleValIleValIleValIle 181
QY 1184 CACGCTAAATCCGCTCTCAAGAAAGAGCTCCGAAAGAGAGGAGAGATATCACCAACCAATC 1243
Db 182 HisAlaLysSerValSerLysLysGlySer---GluGluGlyAspIleThrAsnProIle 200
QY 1244 AACTTGAGAGAGGAGCCGATCTTCTTAACAACCTTTGGGAAGTATTGAGGTGAAG 1303
Db 201 AsnLeuArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLys 220
QY 1304 CCAGACAGAAAGACCCCGAGCTTCAGGACCTGGACATGCTCACCTGTGTAGATC 1363
Db 221 ProAspLysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIle 240
QY 1364 AAAGAGAGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCTCGTC 1423
Db 241 LysGluGlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValVal 260
QY 1424 AACAAAGAACTGGAACCTTGAACCTGCTGTGTAAGAAAGAGACAAACAGAGGGA 1483
Db 261 AsnLysGlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnArgGly 280
QY 1484 CGCGGGAAGAGAGGACCAAGACGAGAGAGAGGAGGAGGAGTAACAGAGAGTGCCT 1543
Db 281 ArgArgGluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArg 300
QY 1544 AGGTACACAGCGAGGTGAAGAAAGCGATGTTCATCATGCCAGCAGCTCATCAGTA 1603
Db 301 ArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProVal 320
QY 1604 GCCATCAACGCTTCCTCCGAACTCCATCTGCTTGGCTTCGATATCAACGCTGAAACAC 1663
Db 321 AlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsn 340
QY 1664 CACAGATCTCTTCGAGGTATAGGACATGTATAGACCATAGACATAGAGAGCAAGCG 1723
Db 341 HisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAla 360
QY 1724 AAGGATTTAGCATTCCTCGGTGAAACAAAGTTGAGAACTCATCAAAAACAGAG 1783
Db 361 LysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLys 380
QY 1784 GAATCTCACTTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCGCTCGTCTGAG 1843
Db 381 GluSerHisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerProGlu 400
QY 1844 AAGAGTCTCTGAGAGAGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1903
Db 401 LysGluSerProGluLysGluAspGlnGluGluGluGluGluGluGluGluGluGluGlu 420
QY 1904 CTTTCAATTTGAAGCTTTTAAAC 1927
Db 421 LeuSerIleLeuLysAlaPheAsn 428

RESULT 6
Q6PSUS ARAHY
ID Q6PSUS ARAHY PRELIMINARY; PRT; 299 AA.
AC Q6PSUS
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Conarachin (Fragment).
OS Arachis hypogaea (peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids 1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;

NR NUCLEOTIDE SEQUENCE.
RP Yan Y.S., Lin X.D., Zhang Y.S., Wang L., Wu K., Huang S.Z.;
RT "Isolation of peanut genes encoding arachins and conglutins by
RT expressed sequence tags."; 2005.
DR EMBL; AY581850; AAT00595.1; -; mRNA.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin_region.
DR Pfam; PF00190; Cupin_1; 1.
FT NON_TER 1
SQ SEQUENCE 299 AA; 33604 MW; 594C3AB0C2FD49EB CRC64;
Alignment Scores:
Pred. No.: 2,95e-90 Length: 299
Score: 1495.50 Matches: 296
Percent Similarity: 97.4% Conservative: 2
Best Local Similarity: 96.7% Mismatches: 1
Query Match: 41.7% Indels: 7
Gaps: 2
US-10-728-323-1 (1-2032) x Q6PSUS_ARAHY (1-299)
QY 1010 TTGAGGCGCCCTTCAATGCGGAATTCATGAGATACGAGGGTCTGTTAGAGAGAT 1069
Db 1 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn 20
QY 1070 GCAGAGGTGAGCAAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1129
Db 21 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTyrSerThrArgSerGluAsn 40
QY 1130 AATGAAGAGGTGATAGTCAAAAGTGTCAAGAGGACAGCTTGAAGAACTTACTTAAGCACGCT 1189
Db 41 AsnGluGlyValIleValLysValSerLysGluHisValGluLeuLeuThrLysHisAla 60
QY 1190 AAATCCGCTCTCAAGAAAGGCTCCGAAAGAGGAGAGATATCACCAACCAATCAACTTG 1249
Db 61 LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu 80
QY 1250 AGAGAGGCGGAGCCGATCTTCTTAACAACCTTGGGAAGTATTTGAGGTGAAGCCAGAC 1309
Db 81 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 100
QY 1310 AAGAAAGAACCCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369
Db 101 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluLysGlu 120
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCGTCTGTCACAAA 1429
Db 121 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 140
QY 1430 GGAATCGAAACCTTGAACCTGCTGCTGTAAGAAAGACCAACAGAGGGGAGCGCG 1489
Db 141 GlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnArgGlyArgArg 160
QY 1490 GAAGAAGAGGAGGAGCAAGAGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1549
Db 161 GluGluGluGluAspGluAspGlnGluGluGluGlySerAsnArgGluValArgArgTyr 180
QY 1550 ACAGCGAGGTGAAGAAAGGCGATGTTTCATCATGCCAGAGCTCATCCATGAGCATC 1609
Db 181 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 200
QY 1610 AACGCTTCTCCGAACCTTCATCTGCTTGGCTTCGATATCAAGCTTGAAGAAACACACACA 1669
Db 201 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 220
QY 1670 ATCTCTCTTGCAGGTGATAAGGACAAATGTGTATAGACCATAGAGAGCAAGCAAGGAGAT 1729
Db 221 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 240
QY 1730 TTGACATTCCTCGGTGCGGTGAAACAAGTTGAGAAAGCTCATCAAAAACAGAGGATCT 1789

DR Pfam; PF00190; Cupin_1; 2.
SQ SEQUENCE 621 AA; 72247 MW; 0DF4B6B37337B7F CRC64;

Alignment Scores:

Pred. No.: 2.15e-77 Length: 621
Score: 1302.00 Matches: 286
Percent Similarity: 61.0% Conservative: 121
Best Local Similarity: 42.9% Mismatches: 170
Query Match: 36.3% Indels: 90
DB: 2 Gaps: 19

US-10-728-323-1 (1-2032) x Q7XXT2_SOYBN (1-621)

QY 50 ATGAGAGGAGGAGGTTTCTCCACTGATGCTGTTCTAGGATCCTTGTCTGGCTTCAGTTT 109
DB 2 MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyValPheLeuAlaSerVal 19
QY 110 TCTGCAACGATGCCAAGTCATCCCTTACCAGAGAAACAGAGAACCCCTCGGCCAG 169
DB 20 SerValSerPheGlyIleAla-----TyrTrpGluLys-----GlnAsnProSerHisAsn 36
QY 170 AGTGTCTCCAGAGTTGTCAACAGGACCGGATGCTTGAAGCAAAAGGCATCGAGTCT 229
DB 37 LysCysLeuArgSerCysAsnSerGluLysAspSerTyrArgAsnGlnAlaCysHisAla 56
QY 230 CGCTGCACC-----AAGCTCGAGTATGATCCTCGTTGT-----GTCATGAT 271
DB 57 ArgCysAsnLeuLeuLysValGluGluGluGluCysGluGluGlyGlnIleProArg 76
QY 272 CTTCTGA---GGACACTGGCACCCACCAACCAACGTTCCCTCCAGGGAGCGGACACGT 328
DB 77 ProArgProGlnHisProGluArgGluArgGlnGlnHisGlyLysGluGluAspGlu 96
QY 329 GGCGGCAACCCGGAGCTACGATGATGACCGC---CGTCAACCCCGAGAGAGGAA--- 382
DB 97 GlyGluGlnProArgProPheProPheProArgProArgGlnProArgGlnGluGlu 116
QY 383 -----GGAGCGCCGATCG 394
DB 117 HisGluGlnLysGluHisGluTrpHisArgLysGluGluLysHisGlyLysGly 136
QY 395 GGA-----CCAGCTGGACCGAGGAGCGTGA 421
DB 137 SerGluGluGluGlnAspGluArgGluHisProArgProHisGlnProHisGlnLysGlu 156
QY 422 AGAGAGAA---GACTGGAGACACCAACAGAGAGATTGGAGCGCACCAAGTCATCAGCAG 478
DB 157 GluGluLysHisGluTrpGlnHisLysGlnGluLysHisGlnGlyLysGluSerGluGlu 176
QY 479 CCACGGAAATAAGGCCCGAAGGAGAGAGAGAAACAAG---TGGGGAACACCAAGCT 535
DB 177 GluGluGluAspGlnAspGluGluGluGlnAspLysGluSerGlnGluSerGluGly 196
QY 536 AGCCATGTGAGGAAGAAACATCTCGAAC-----AACCTTTCTACTTCCCGTCA 586
DB 197 SerGluSerGlnArgGluProArgArgHisLysAsnLysAsnProPheHisPheAsnSer 216
QY 587 AGGCGGTTTAGCACCGCTACGGGAACCAACACGCTAGGATCCGGCTCTCGCAGAGGTTT 646
DB 217 LysArgPheGlnThrLeuPheLysAsnGlnTyrGlyHisValArgValLeuGlnArgPhe 236
QY 647 GACCAAGGTCAAGGAGGTTTCAAGATCTCCAGATCCACGCTATTGTGCGATCGAGGCC 706
DB 237 AsnLysArgSerGlnGlnLeuGlnAsnLeuArgAspTyrArgIleLeuGluPheAsnSer 256
QY 707 AAACCTTAACACTTGTCTTCCCAAGCAGCGTGTGATGATCAACATCCCTGTTATCCAG 766
DB 257 LysProAsnThrLeuLeuProHisHisAlaAspAlaAspTyrLeuIleValIleLeu 276
QY 767 CAAGGCAACCCACCGTACCGTAAATGGCAATTAACAGAAAGAGCTTTAATCTTGAC 826
DB 277 AsnGlyThrAlaIleLeuThrLeuValAsnAsnAspAspAspSerTyrAsnLeuGln 296

QY 827 GAGGGCCACTGCACATCCCATCCCGTTCATTTCCTACATCTTGAACCCCATGAC 886
DB 297 SerGlyAspAlaLeuArgValProAlaGlyThrThrTyrValValAsnProAspAsn 316
QY 887 AACCAAGACCTCAGAGTAGCTAAATCTCCATGCCGTTAAACACACCCCGCCAGTTT 946
DB 317 AspGluAsnLeuArgMetIleThrLeuAlaIleProValAsnLysProGlyArgPheGlu 336
QY 947 GATTTCCTCCCGCGCAGCAGCGACCAATCATCTACTTGCAGGCTTTCAGCAGAAT 1006
DB 337 SerPheLeuSerSerThrGlnAlaGlnSerTyrLeuGlnGlyPheSerLysAsn 356
QY 1007 ACGTTGGAGCGCCCTTCAATCGGAATTCATAGATACGGAGGCTGCTGTAGAGAG 1066
DB 357 IleLeuGluAlaSerTyrAspThrLysPheGluGluIleAsnLysValLeuPheGlyArg 376
QY 1067 AATCAGGAGGTGAGCAA---CAGGAGAGCGCAGCGGATGGAGTACTCGGAGTAGT 1123
DB 377 GluGluGlyGlnGlnGlnGlyGluGluArgLeuGln----- 388
QY 1124 GAGAACAAATGAAGGAGTAGTCAAAGTGTCAAAGGACACGTTCAAGAACTTACTAAG 1183
DB 389 -----GluSerValIleValGluIleSerLysLysGlnIleArgGluLeuSerLys 405
QY 1184 CACGCTAAATCCGTCTCAAAGAAAGCTCCGAAGAGAGAGGAGATATCAACCAACCAAT 1243
DB 406 HisAlaLysSerSerSerArgLysThrIleSerSerGlu-----AspLysProPhe 422
QY 1244 AACTTGAGAGAGCGCGCCGATCTTCTTCAACACTTTGGGAAGTATTATTAGGTGAAG 1303
DB 423 AsnLeuArgSerArgAspProIleTyrSerAsnLysLeuGlyLysLeuPheGluIleThr 442
QY 1304 CAGACAAAGAAACCCCGCTTCAGGACCTGGACATCATGCTACCTGTGTAGAGATC 1363
DB 443 ProGlu---LysAsnProGlnLeuArgPheLeuAspValPheLeuSerValValAspMet 461
QY 1364 AAAGAGGAGCTTGTGATGCTCCACACTCAAAGGCCATGGTTATCGTCGCTG 1423
DB 462 AsnGluGlyAlaLeuPheLeuProHisPheAsnSerLysAlaIleValValLeuValIle 481
QY 1424 ACAAGAGNACTGGAAACCTTGAACTCGTGGCTGTGAAGAAAGAGACAAACAGAGGGGA 1483
DB 482 AsnGluGlyGluAlaAsnIleGluLeuValGlyIle---LysGluGlnGlnArgGln 500
QY 1484 CGCGGGAAGAGGAGGACGACGACGAGAGAGAGAGGAGGAGGAGTAAACAGAGAGTCCGT 1543
DB 501 GlnGlnGluGluGlnPro-----LeuGluValArg 510
QY 1544 AGGTACACAGCGAGGTGAAGGAGCGGATGTGTTTCATCATGCCACAGCTCATCCAGTA 1603
DB 511 LysTyrArgAlaGluLeuSerGluGlnAspIlePheValIleProAlaGlyTyrProVal 530
QY 1604 GCATCAAGCTTCTCCGAACCTCCATCTGCTTGGCTTCGGTATCAACCTGNAACCAAC 1663
DB 531 ValValAsnAlaThrSerAspLeuAsnPhePheAlaPheGlyIleAsnAlaGluAsnAsn 550
QY 1664 CACAGATCTCTCTGCGGTGATAGGACAAATGTGATAGACCAGATAGAGAACGCG 1723
DB 551 GlnArgAsnPheLeuAlaGlySerLysAspAsnValIleSerGlnIleProSerGlnVal 570
QY 1724 AAGGATTTAGCATTCCTCGGCTGAACAAGTTGAGAAGCTCATCAAAACACAGAAG 1783
DB 571 GlnGluLeuAlaPheProGlySerAlaLysAspIleGluAsnLeuIleLysSerGlnSer 590
QY 1784 GAATCTCACTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCCGCTCGTCTCTGAG 1843
DB 591 GluSerTyrPheValAspAlaGlnProGln----- 600
QY 1844 AAGAGTCTCTCGAGAAAGAGGATCAAGAGGAGGAGAAACCAAGGAGGAGGCTCCACTC 1903
DB 601 -----GlnLysGluGluGlyAsnLysGlyArgLysGlyProLeu 613
QY 1904 CTTTCAATTTTGAAGGCTTTT 1924


```
QY 1604 GCCATCAACGGCTTCCTCCGAACCTCCATCTGCTTGGCTTCGGTATCAACGCTGAAAAACAC 1663
Db      ::::::::::::::::::::::
531 ValValAsnAlaThrSerAspLeuAsnPhePheAlaPheGlyLeuAsnAlaGluAsnAsn 550
QY 1664 CACAGAACTTCCTTCGACGGTGAATAGGACAATGTGATAGACAGATAGAGAAGCAAGCG 1723
Db      |||
551 GlnArgAsnPheLeuAlaGlySerLysAspAsnValIleSerGlnIleProSerGlnVal 570
QY 1724 AAGGATTTAGCATTCCTCGGTCCGGTGAACAAGTTGAGAACTCATCAAAAACCAAGAG 1783
Db      ::::::::::::::::::::::
571 GlnGluLeuAlaPheProGlySerAlaLysAspIleGluAsnLeuIleLysSerGlnSer 590
QY 1784 GAATCTCACTTTGTAGTCTCGTCTCAATCTCAATCTCAATCTCCGTCGTCCTCGAG 1843
Db      |||
591 GluSerTyrPheValAspAlaGlnProGln----- 600
QY 1844 AAGAGTCTCTGAGAAAGAGATCAAGAGGAGGAAACCAAGGAGGAGGAGGTGCATCTC 1903
Db      ::::::::::::::
601 -----GlnLysGluGluGlyAsnLysGlyArgLysGlyProLeu 613
QY 1904 CTTTCAATTTGAAGGCTTTT 1924
Db      |||
614 SerSerIleLeuArgAlaPhe 620
```

RESULT 10

```
Q6EBC1_LUPAL
ID Q6EBC1_LUPAL PRELIMINARY; PRT; 533 AA.
AC Q6EBC1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Beta-conglutin.
OS Lupinus albus (White lupin).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosoids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.
RX NCBI_TaxID=3870;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RA Monteiro S.A., Freitas R.M., Teixeira A.N., Ferreira R.B.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY500372; AAS97865.1; -; mRNA.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin_region.
DR Pfam; PF00190; Cupin_1; 2.
SQ SEQUENCE 533 AA; 62129 MW; F17B49A6F2C4B9EC CRC64;
```

Alignment Scores:

```
Pred. No.: 1.29e-76 Length: 533
Score: 1290.00 Matches: 273
Percent Similarity: 62.9% Conservative: 117
Best Local Similarity: 44.0% Mismatches: 132
Query Match: 36.0% Indels: 98
DB: Gaps: 17
```

US-10-728-323-1 (1-2032) x Q6EBC1_LUPAL (1-533)

```
QY 44 GGAGCAATCAGAGGGAGGGTTTCTCCATCGATGCTTCTCCTAGGATCCTTGTCTCGTGGCT 103
Db      |||
2 GlyLysMetArgValArgPheProThrLeuValLeuValLeuGlyValPheLeuMet 21
QY 104 TCAGTTTCTGCACGATGCCAAGTCATCACCTTACCAGAAAGAAACAGAGAACCCCTGC 163
Db      ::::::::::::::
22 AlaValSerIleGlyIleAla-----TyrGlyGluLysAsp----- 33
QY 164 GCCCAGAGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATGC 223
Db      ::
34 -----ValLeuLysSerHisGluArgProGluGluArgGluGluGln----- 47
QY 224 GAGTCTCGTCGACCAAGCTCGAGTATGATCTCGTTGTGTCTATGATCCTCGAGACAC 283
Db      |||
```

```
Db      -----GluTrpGlnProArg----- 52
QY 284 ACTGSCACCACCAACCAACAGTTCCCTCCAGGGAGCGACACGTCGCGCCCAACCCGGA 343
Db      |||
53 -----ArgGlnArgPro----- 56
QY 344 GACTACGATGATACCGCGCTCAACCCGGAAGAGAGAGGAGCGCATGGGACCAGCT 403
Db      |||
57 -----GlnSerArgArgGluArgGluGlnGluGlnGlnGlnGlnGln 73
QY 404 GGACCGAGGAGCGTGAAGAGAAAGACATCGGACACACCAAGAGAGATTTGGAGGCCA 463
Db      |||
74 TyrProArgArgGlnSerGlyTyrGlu-----ArgArg 84
QY 464 CCAAGTCATCAGCAGCCAGGAAATAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGTGG 523
Db      |||
85 GlnTyrHisGlu-----ArgSerGluGlnArgGluArgGluGlnGluGlnGlnGln 102
QY 524 GGAACACAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACCAACCTTTTCTACTCCCG 583
Db      |||
103 GlySerPro---SerTyrSerArgArgGln-----ArgAsnProTyrHisPheSer 118
QY 584 TCAAGCGGTTTAGCACCGCTACCGGAACCAAAACGGTAGGATCCGGTCTCTGCAGAGG 643
Db      |||
119 SerGlnArgPheGlnThrLeuTyrLysAsnArgAsnGlyLysIleArgValLeuGluArg 138
QY 644 TTTGACCAAGGTCAGGCGATTTCCAGAACTCTCCAGAACTCCACGATTTGTGCAGATCGAG 703
Db      |||
139 PheAspGlnArgThrAsnArgLeuGluAsnLeuGlnAsnTyrArgIleValGluPheGln 158
QY 704 GCCAAACCTTAACACTCTTGTCTTCCCAAGCACGCTGATGTGTGATAACATCTCTGTATC 763
Db      |||
159 SerLysProAsnThrLeuIleLeuProLysHisSerAspAlaAspTyrValLeuValVal 178
QY 764 CAGCAAGGGCAAGCCACGTCACCTAGCAATGCAATGCAATACAGAAAGAGCTTTAATCTT 823
Db      |||
179 LeuAsnGlyArgAlaThrIleThrIleValAsnProAspArgArgGlnAlaTyrAsnLeu 198
QY 824 GACGAGGGCCATGCACCTCAGAACTCCATCCGTTTTCATTCTCATCTTTGAACGCCAT 883
Db      |||
199 GluTyrGlyAspAlaLeuArgIleProAlaGlySerThrSerTyrIleLeuAsnProAsp 218
QY 884 GACACCAAGAACCTCAGAGTAGTAAATCTCCATGCCCGTTTAAACACACCCCGCCAGTTT 943
Db      |||
219 AspAsnGlnLysLeuArgValValLysLeuAlaIleProIleAsnAsnProGlyTyrPhe 238
QY 944 GAGGATTTCTCCCGCGAGCAGCGACACCAATCATCTCTACTTCGAGGGCTTCAGCAGG 1003
Db      |||
239 TyrAspPheTyrProSerSerThrLysAspGlnSerTyrPheSerGlyPheSerArg 258
QY 1004 AATCGTTTGGAGGCGCTTCAATCGGAAATTCATAGATACGAGGAGTGTGTTA--- 1060
Db      |||
259 AsnThrLeuGluAlaThrPheAsnThrArgTyrGluGluIleGlnArgIleLeuGly 278
QY 1061 ---GAAGAGAATTCAGAGGAGTGCAGCAAGAGAGAGGCGAGGAGGAGTGGAGTACTCGG 1117
Db      |||
279 AsnGluAspGluGlnGluTyrGluGluGlnArgArgGlyGlnGluGln----- 294
QY 1118 AGTAGTGACAATGAAGGAGTGATAGTCAAGTGTCAAGGAGCAGCGTGAAGAACTT 1177
Db      |||
295 ---SerAspGlnAspGluGlyValIleValIleValSerLysLysGlnIleGlnLysLeu 313
QY 1178 ACTAGCAGCTAAATCCGTCTCAAGAAAGAGGCTCCGAAAGAGAGAGGAGATATACCAAC 1237
Db      |||
314 ThrLysHisAlaGlnSerSerSer-----GlyLysAspLysProSerAsp---SerGly 330
QY 1238 CCAATCAACTTGAGAGAGGCGAGCCGATCTTCTTAAACACTTTGGGAGGTTATTGAG 1297
Db      |||
331 ProPheAsnLeuArgSerAsnGluProIleTyrSerAsnLysTyrGlyAsnPheTyrGlu 350
QY 1298 GTGAGGCCAGACAGACAGAAAGACCCAGCTTCAGGACCTGGACATGATGCTACCTGTGTA 1357
Db      |||
351 IleThrProAsp---ArgAsnProGlnValGlnAspLeuAsnIleSerLeuThrTyrIle 369
```

QY 1358 GAGATCAAGAGAGAGCTTTGATGCTCCACACTTCAACTCAAGGCGATGTTATCGTC 1417
 Db LysileAsnGluGlyAlaLeuLeuProHisIleAsnSerLysAlaIleTyrValVal 389
 QY 1418 GTCTCAACAAAGAACTGGAACCTTGAACCTGCTGCTGTAAGAAAGAGCAACACAG 1477
 Db ValValAspGluGlyGluGlyAsnTyrGluLeuValGlyIleArgAspGlnGlnArgGln 409
 QY 1478 AGGGGACGGCGGAAG 1537
 Db 410 -----GlnAspGluGlnGluGluGlyGlu-----GluGlu 419
 QY 1538 GTCCGTAGGTACACAGGAGTTGAAGAGAGGCGATGTTTCATCATGCGCAGAGCTCAT 1597
 Db 420 ValIleArgTyrSerAlaArgLeuSerGluGlyAspIlePheValIleProAlaGlyTyr 439
 QY 1598 CCAGTAGCCATCAACGCTTCCTCCGAACCTCCATCTGCTGCTTCGGTATCAACGCTGAA 1657
 Db 440 ProIleSerIleAsnAlaSerSerAsnLeuArgLeuGlyPheGlyIleAsnAlaAsp 459
 QY 1658 AACCAACACAGAAATCTTCTTGCAGGTGATAAGGACAAATGTATAGACCAAGATAGACAAG 1717
 Db 460 GluAsnGlnArgAsnPheLeuAlaGlySerLysAspAsnValIleArgGlnLeuAspArg 479
 QY 1718 CAACGAGAGATTTAGCATTCCTGCGTGGGTGAACAAGTTGAGAAGCTCATCAAAAAC 1777
 Db 480 AlaValAsnGluLeuThrPheProGlySerAlaGluAspIleGluArgLeuIleLysAsn 499
 QY 1778 CAGAAGAACTCACTTTGTGAGTGTGCTGCTCAATCTCAATCTCAATCTCGTCTCT 1837
 Db 500 GlnGlnGlnSerTyrPheAlaAsnGlnGlnProGlnGlnGln----- 515
 QY 1838 CCTGAGAAAGAGTCTCTGAGAAAGAGATCAAGAGGAGGAGAAACCAAGAGGAGGAGGT 1897
 Db 516 -----GlnGlnSerGluGlyArgArgGly 525

RESULT 11

Q4LERS_SOYBN PRELIMINARY; PRT; 604 AA.
 AC Q4LERS;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Beta-conglycinin alpha subunit (fragment).
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 OC Glycine.
 OX NCBI_TaxID=3847;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Taubokura Y., Hajika M., Harada K.;
 RT "Molecular characterization of a beta-conglycinin Deficient Soybean.";
 RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB197785; BAB02727.1; -; mRNA.
 FT NON TER 1
 SQ SEQUENCE 604 AA; 70064 MW; C2805CCCB4C48507 CRC64;

Alignment Scores:

Pred. No.: 2.44e-76 Length: 604
 Score: 1286.00 Matches: 286
 Percent Similarity: 59.5% Conservative: 108
 Best Local Similarity: 43.2% Mismatches: 172
 Query Match: 35.9% Indels: 96
 DB: 2 Gaps: 18

US-10-728-323-1 (1-2032) x Q4LERS_SOYBN (1-604)

QY 50 ATGAGAGGAGGGTGTCTCCACTGATGCTGTGTAGGATCCTGTGCTGGTTCAGTT 109
 Db LysileAsnGluGlyAlaLeuLeuProHisIleAsnSerLysAlaIleTyrValVal 389
 1 MetArgAlaArgPhePro-----LeuLeuLeuGlyLeuValPheLeuAlaSerVal 18

QY 110 TCTGCAACGCATGCCAGTCATCCTTTACAGAGAAACACAGAACCCCTGCGCCAG 159
 Db SerValSerPheGlyIleAla-----TyrTrpGluLys---GluAsnProLysHisAsn 35
 QY 170 AGGTGCTCCAGAGTTGTCAACAGAGAACCGGATGACTTGAAGCAAAAGCATGCGAGTCT 229
 Db LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 55
 QY 230 CGCTGCACCAAGCTCGAGTATGAT-----CCTCGTTGT 262
 Db 56 ArgCysAsnLeuLeuLysValGluLysGluGluCysGluGluGlyGluIleProArg--- 74
 QY 263 GTCTATGATCTCTCGAGGACACACTGGCACCACCAACCAACGTTCCCTCCAGGGAGCG 322
 Db 75 -----ProArgProArg-----ProGlnHisProGlu 83
 QY 323 ACACGTGGCGGCAACCCCGAGACTACGATGATGAC----- 358
 Db 84 ArgGluProGlnGlnProGlyGluLysGluAspGluAspGluGlnProArgProIle 103
 QY 359 -----CGCGTCAACCCCGAAGAGAG-----GAAGAGGCCGA 391
 Db 104 ProPheProArgProGlnProArgGlnGluGluHisGluGlnSerGluGluGlnGlu 123
 QY 392 TGG-----GGACCAGCTGGACCGGAGCGTGAAAGAGAGAA 430
 Db 124 TrpProArgLysGluGluLysArgGlyGluLysGlySerGluGluGluAspGlu 143
 QY 431 GACTGGAGACAACCAAGAGAAGAT-----TGAGGCGCACCAAGTCAATCAGCACCCAG 484
 Db 144 AspGluGluGlnAspGluArgGlnPheProPheProArgProProHisGlnLysGluGlu 163
 QY 485 AAAATAAGCCCGCAAG 544
 Db 164 ArgLysGlnGluGluAspGluAspGluGluGlnGlnGlnSerGluGluSerGluAsp 183
 QY 545 AGGGAAGAAACATCTCGGAAC-----AACCCCTTCTACTTCCCGTCAAGCGCGTTAGC 598
 Db 184 SerGluLeuArgArgHisAsnAsnLysAsnProPheLeuPheGlySerAsnArgPheGlu 203
 QY 599 ACCCGCTAGCGGAACCAAAACCGTAGGATCCCGGCTCTCGCAGAGGTTTGACCAAGGTCA 658
 Db 204 ThrLeuPheLysAsnGlnTyrGlyArgIleArgValLeuGlnArgPheAsnGlnArgSer 223
 QY 659 AGGCAGTTTCAGAAATCTCCAGATCACCGTATTGTGCAGATCGAGGCCAACCTAACACT 718
 Db 224 ProGlnLeuGlnAsnLeuArgAspTyrArgIleLeuGluPheAsnSerLysProAsnThr 243
 QY 719 CTGTGTTCTTCCCAAGCACGCTGATGCTGATAACATCTTGTATTCCAGCAAGGCGCAAGCC 778
 Db 244 LeuLeuLeuProAsnHisAlaAspAlaAspTyrLeuIleValIleLeuAsnGlyThrAla 263
 QY 779 ACCGTGACCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAGGCCATGCA 838
 Db 264 IleLeuSerLeuValAsnAsnAspArgAspSerTyrArgLeuGlnSerGlyAspAla 283
 QY 839 CTCAGAAATCCCATCCCGTTTCATTTCTTACATCTTGAACCGCCATCACACACCAACCTC 898
 Db 284 LeuArgValProSerGlyThrTyrTyrValValAsnProAspAsnAsnGluAsnLeu 303
 QY 899 AGAGTAGTAAATCTCCATGCCCGCTTAACACACCCCGGCGAGTTTCTAGGATTTCTTCCCG 958
 Db 304 ArgLeuIleThrLeuAlaIleProValAsnLysProGlyArgPheGluSerPhePheLeu 323
 QY 959 GCGAGCAGCGAGACCAATCATCTACTTGCAGGGCTTCAGCAGGAATACGTTGGAGGCC 1018
 Db 324 SerSerThrGluAlaGlnSerTyrLeuGlnGlyPheSerArgAsnIleLeuGluAla 343
 QY 1019 GCCTTCATCGGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1078
 Db 344 SerTyrAspThrLysPheGluGluIleAsnLysValLeuPheSerArgGluGluGly--- 362

QY	623	AGGATCCGGTCTCTGAGAGGTTTGACCAAGGTCAGGTCCTCCAGAAT	682
Db	229	HisValArgValLeuGlnArgPheAsnLysArgSerGlnGlnLeuGlnAsnLeuArgAsp	248
QY	683	CACCGTATTGTGCAGATCGAGGCGCAACCTACACTCTTTGTTCTTCCCAAGCAGCGTGAT	742
Db	249	TyrArgIleLeuGluPheAsnSerLysProAsnThrLeuLeuLeuProHisAlaAsp	268
QY	743	GCTGATAACATCTTGTATCCAGCAAGGCGCAAGCCCGTACCGTAGCAATGGCAAT	802
Db	269	AlaAspTyrLeuIleValIleLeuAsnGlyThrAlaIleLeuThrLeuValAsnAsp	288
QY	803	AACAGAAAGAGCTTTAATCTTGACGAGGCGCCATGCACTCAGAAATCCCGTTCATT	862
Db	289	AspArgAspSerTyrAsnLeuGlnSerGlyAspAlaLeuArgValProAlaGlyThrThr	308
QY	863	TCCTACATCTTGAAACCGCCATGACACACCAAGACCTCAGAGTAGCTAAATCTCCATGCC	922
Db	309	TyrTyrValValAsnProAspAsnAspGluAsnLeuArgMetIleThrLeuAlaIlePro	328
QY	923	GTTAACACACCCCGCAGTTTGAGGATTTCTCCCGCGAGCAGCGAGCAATCATCC	982
Db	329	ValAsnLysProGlyArgPheGluSerPhePheLeuSerSerThrGlnAlaGlnSer	348
QY	983	TACTTGAGGGCTTCAGCAGGAATACGTTGGAGGCGCCTTCAATGCGGAATTCATAGAG	1042
Db	349	TyrLeuGlnGlyPheSerLysAsnIleLeuGluAlaSerTyrAspThrLysPheGluGlu	368
QY	1043	ATACGAGGGTCTGTTAGAGAAGATGCGAGGAGTGAGCAA--GAGGAGAGGCGCAG	1099
Db	369	IleAsnLysValLeuPheGlyArgGluGluGlyGlnGlnGlyGluGluArgLeuGln	388
QY	1100	AGGCGATGGAGTACTCGAGTAGACAACATGAAGGAGTGATAGTCAAAGTGTCAAAG	1159
Db	389	-----GluSerValIleValGluIleSerLys	397
QY	1160	GAGCAGCTTGAAGACTTACTAAGCAGCGTAAATCCGTCTCAAGAAAGGCTCCGAGAA	1219
Db	398	LysGlnIleArgLeuSerLysArgAlaLysSerSerArgLysThrIleSerSer	417
QY	1220	GAGGAGATATACCAACCAATCAACTTGAGAGAAGGCGAGCCGATCTTTTAAACAAC	1279
Db	418	Glu-----AspLysProPheAsnLeuArgSerArgAspProIleTyrSerAsnLys	434
QY	1280	TTTGGGAAGTTATTGAGGTGAAGCCAGACAAGAACCCCGAGCTTCAGGACCTGGAC	1339
Db	435	LeuGlyLysLeuPheGluIleThrProGlu--LysAsnProGlnLeuArgAspLeuAsp	453
QY	1340	ATGATGTCACCTGTGTAGAGATCAAGAAGGAGCTTTGATGCTCCACACTTCAACTCA	1399
Db	454	ValPheLeuSerValValAspMetAsnGluGlyAlaLeuPheLeuProHisPheAsnSer	473
QY	1400	AAGGCCATGTTATCGTCGTCGTCAACAAAGGAACCTTGAACTCTGTCGCTGTA	1459
Db	474	LysAlaIleValValLeuValIleAsnGluGlyGluAlaAsnIleGluLeuValGlyIle	493
QY	1460	AGAAAAGAGCAACAACAGAGGCGCGCGGGAAGAAGAGGAGGACGAAAGCAAGAGAG	1519
Db	494	---LysGluGlnGlnArgGlnGlnGlnGluGlnPro-----	506
QY	1520	GAGGAAGTAAACAGAGGTCGTAGGTACACAGCGAGGTTGAAGNAGCGGATGTGTC	1579
Db	507	-----LeuGluValArgLysTyrArgAlaGluLeuSerGluGlnAspIlePhe	522
QY	1580	ATCATGCCAGAGCTCATCAGTAGGCATCAACGCTTCCTCCGAACTCCATCTGCTGGC	1639
Db	523	ValIleProAlaGlyTyrProValValValAsnAlaThrSerAspLeuAsnPhePheAla	542
QY	1640	TTCCGTATCAACGCTGAAAAACAACACAGAAATCTTCTTCGAGGTGATAGCAATGTG	1699
Db	543	PheGlyIleAsnAlaGluAsnAsnGlnArgAsnPheLeuAlaGlySerLysAspAsnVal	562

QY	1700	ATAGACCAGATAGAGCAAGCAAGGATTTAGCATTCCTCGGTCCGGTGAAACAAGTT	1759
Db	563	IleSerGlnIleProSerGlnValGlnGluLeuAlaPheLeuGlySerAlaLysAspIle	582
QY	1760	GAGAAGCTCATCAAAACACAGAAAGTAATCTCACCTTTGTGAGTGTCTCAATCTCAA	1819
Db	583	GluAsnLeuIleLysSerGlnSerGluSerTyrPheValAspAlaGlnProGln-----	600
QY	1820	TCTCAATCTCCGTCTCTCTGAGAAAGAGTCTCTCTGAGAAAGAGGATCAAGAGGAGAA	1879
Db	601	-----GlnLysGluGluGly	605
QY	1880	AACCAAGGAGGAGGTCCTCTCTTCAATTTTGAAGGCTTTT	1924
Db	606	AsnLysGlyArgLysGlyProLeuSerSerIleLeuArgAlaPhe	620

RESULT 13

Q94LX2_SOYBN

IDQ94LX2_SOYBNPRELIMINARY;PRT;605AA.

ACQ94LX2;

DT01-DEC-2001(TREMBLrel.19,Created)

DT01-DEC-2001(TREMBLrel.19,Lastsequenceupdate)

DT01-JUN-2003(TREMBLrel.24,Lastannotationupdate)

DEBeta-conglycinin alphasubunit.

OSGlycine max (Soybean).

OCEukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

OC Glycine.

OC NCBI_TaxID=3847;

RN[1]

RPNUCLEOTIDESEQUENCE.

RX MEDLINE=21327318; PubMed=11434464; DOI=10.1266/ggs.76.99;

RAYoshino M., Kanazawa A., Tsutsumi K., Nakamura I., Shimamoto Y.;

RT "Structure and characterization of the gene encoding alphasubunit of

RL soybean beta-conglycinin.";

RL Genes Genet. Syst. 76:99-105(2001).

RN[2]

RPNUCLEOTIDESEQUENCE.

RX PubMed=1731988;

RALelievre J.M., Dickinson C.D., Dickinson L.A., Nielsen N.C.;

RT "Synthesis and assembly of soybean beta-conglycinin in vitro.";

RL Plant Mol. Biol. 18:259-274(1992).

DR EMBL; AB051865; BAB56161.1; -; Genomic_DNA.

DR PIR; S20007; S20007.

DR HSSP; P25974; 11PJ.

DR SMR; Q94LX2; 194-581.

DR GO; GO:0045735; F:nutrient reservoir activity; IEA.

DR InterPro; IPR006045; Cupin.

DR InterPro; IPR007113; Cupin_region.

DR Pfam; PF00190; Cupin_1; 2.

SQSEQUENCE605AA; 70306MW; 8ACE6F8532662984CRC64;

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

4,49e-76

1282.00

59.4%

43.1%

35.8%

2

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

605

285

108

173

96

18

US-10-728-323-1 (1-2032) x Q94LX2_SOYBN (1-605)

QY	50	ATGAGAGGAGGTTTCTCCACTGATGCTGTCTAGGATCTTGTCTCGGTCGCTCAGTT	109
Db	2	MetArgAlaArgPhePro-----LeuLeuLeuGlyLeuValPheLeuAlaSerVal	19
QY	110	TCTGCAACGATCCCAAGTCATCACCTTACCAGAAGAAACAGAGAACCCCTCGCCACG	169
Db	20	SerValSerPheGlyIleAla-----TyrTrpGluLys---GluAsnProLysHisAsn	36
QY	170	AGGTGCTCCAGAGTTGTCAACAGGAAACCGGATGACTTGAAGCAAAAGGATCGAGTCT	229

Db 37 LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 56
QY 230 CGCTGCACCAAGCTCCAGTATGAT-----CCTCGTTGT 262
Db 57 ArgCysAsnLeuLeuLysValGluLysGluCysGluGluGlyGluLeuProArg--- 75
QY 263 GTCTATGATCCTCGAGGACACACTGGCACACCAACCAACGTTCCCTCAGGGGAGCGG 322
Db 76 -----ProArgProArg-----ProGlnHisProGlu 84
QY 323 ACACGTGGCGCCCAACCCGAGACTACGATGATGAC----- 358
Db 85 ArgGluProGlnGlnProGlyGluLysGluGluAspGluAspGluProArgProIle 104
QY 359 -----CGCGTCAACCCCGAGAGAG-----GAAGAGGCCGA 391
Db 105 ProPheProArgProGlnProArgGlnGluGluHisGluGlnArgGluGlnGlu 124
QY 392 TGG-----GGACCAAGCTGGACCGAGGAGCGTGAAGAGAGAA 430
Db 125 TrpProArgLysGluGluLysArgGlyGluLysGlySerGluGluGluAspGlu 144
QY 431 GACTGGAGACCAACAGAGAGAT-----TGGAGCGGACCAAGTCAACGAGCCACCG 484
Db 145 AspGluGlnAspGluArgGlnPheProPheProArgProProHisGlnLysGluGlu 164
QY 485 AATAAAGGCCCGAAG 544
Db 165 ArgLysGlnGluGluAspGluGluGlnGlnArgGluSerGluGluSerGluAsp 184
QY 545 AGGGAA-----GAAACATCTCGAAACACCTTTCTACTTCCCGTCAAGCGGTTTAGC 598
Db 185 SerGluLeuArgArgHisLysAsnLysAsnProPheLeuPheGlySerAsnArgPheGlu 204
QY 599 ACCCGTAGCGGACCAACACCGTAGATCCGGTCTCGAGAGGTTTACCAAGAGTCA 658
Db 205 ThrLeuPheLysAsnGlnTyrGlyArgIleArgValLeuGlnArgPheAsnGlnArgSer 224
QY 659 AGCGAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGCCAAACCTAACACT 718
Db 225 ProGlnLeuGlnAsnLeuArgAspTyrArgIleLeuGluPheAsnSerLysProAsnThr 244
QY 719 CTTGTTCTCCCAAGCAACGCTGATGTATACATCTTGTGTTATCCAGCAAGGCGAAGCC 778
Db 245 LeuLeuLeuProAsnHisAlaAspAlaAspTyrLeuIleValIleLeuAsnGlyThrAla 264
QY 779 ACCGTACCGTAGCAATGCGCAATACAGAAAGAGCTTAACTTCAGAGGCCCATGCA 838
Db 265 IleLeuSerLeuValAsnAsnAspAspArgAspSerTyrArgLeuGlnSerGlyAspAla 284
QY 839 CTGAGAACTCCATCCCGTTTCATTCTTACATCTTCAACCGCCATGACAAACAGAACCTC 898
Db 285 LeuArgValProSerGlyThrThrTyrTyrValValAsnProAspAsnAsnGluAsnLeu 304
QY 899 AGATGAGTAAATCTCCATCCCGTTAAACACACCCGCGCAGTTGAGGATTTCTTCCCG 958
Db 305 ArgLeuIleThrLeuAlaIleProValAsnLysProGlyArgPheGluSerPhePheLeu 324
QY 959 GCGAGCAGCGGACCAATCTCTACTTCAGCGGCTTCAGCAGGAATAGTTGGAGGCC 1018
Db 325 SerSerThrGluAlaGlnSerTyrLeuGlnGlyPheSerArgAsnIleLeuGluAla 344
QY 1019 GCCTTCAATCGGAAATTCATAGATACGAGGAGTCTGTAGAGAGAATCGAGAGGT 1078
Db 345 SerTyrAspThrLysPheGluGluIleAsnLysValLeuPheSerArgGluGluGly--- 363
QY 1079 GAGCAAGAGGAGAGCGGAGCGGAGTACTCGGAGTAGTGAGAACAAATGAAGGA 1138
Db 364 -----GlnGlnGlnGlyGluGlnArgLeu-----GlnGluSer 374
QY 1139 GTGATAGTCAAGTGTCAAGAGGACGTTGAGAACTTACTTAAGCACCGCTAAATCCGTC 1198
Db 375 ValIleValIleSerLysGluGlnIleArgAlaLeuSerLysArgAlaLysSerSer 394

QY 1199 TCAAGAAAGAGGCTCCGAAAGAGAGAGGAGATATCACCAACCCCAATCAACTTGAGAGAAGC 1258
Db 395 SerArgLysThrIleSerSerGlu-----AspLysProPheAsnLeuArgSerArg 411
QY 1259 GAGCCCGATCTTTCTAAACAACCTTTGGGAAGTTATTGGTGAAGCCACAGAGAAGAAC 1318
Db 412 AspProIleTyrSerAsnLysLeuGlyLysPhePheGluIleThrProGlu---LysAsn 430
QY 1319 CCCAGCTTCAGACCTGGACATGATCTCACCTGTGTAGAGATCAAGAAGAGAGCTTTG 1378
Db 431 ProGlnLeuArgAspLeuAspIlePheLeuSerIleValAspMetAsnGluGlyAlaLeu 450
QY 1379 ATGCTCCCACTTCAACTCAAAAGCCATGTTATCGTCTCGTCAACAAAGAACTGGA 1438
Db 451 LeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGlyAspAla 470
QY 1439 AACCTTGAACCTCGTGGCTGTAAGAAAGACCAACAGAGGGGACGGGGGAGAGAG 1498
Db 471 AsnIleGluLeuValGlyLeu---LysGluGlnGlnGlnGluGlnGlnGluGln 489
QY 1499 GAGGACGAAGACGAAGAGAGGAGGAGTAACAGAGAGGTGCTAGGTACACAGCGAGG 1558
Db 490 Pro-----LeuGluValArgLysTyrArgAlaGlu 499
QY 1559 TTGAAGGAAGCGATGTGTTTCATCATCCAGCAGCTCATCCAGTAGCCATCAACGCTTC 1618
Db 500 LeuSerGluGlnAspIlePheValIleProAlaGlyTyrProValValValAsnAlaThr 519
QY 1619 TCGGAATCTCATCTCGTTCGGTTCATCAACGCTGAAAGGAGGAGGAGGAGGAGGAG 1678
Db 520 SerAsnLeuAsnPhePheAlaIleGlyIleAsnAlaGluAsnGlnArgAsnPheLeu 539
QY 1679 CGAGGTGATAAGGACAATGTGATAGACAGATAGAGAACAAGCAGGAGGATTTAGCATTC 1738
Db 540 AlaGlySerGlnAsnValIleSerGlnIleProSerGlnValGlnGluLeuAlaPhe 559
QY 1739 CTTGGTTCGGGTGAACAAAGTTGAGAGGCTCATCAAAAACAGAGAAGTCTCACTTTGTG 1798
Db 560 ProGlySerAlaGlnAlaValGluLysLeuLeuLysAsnGlnArgGluSerTyrPheVal 579
QY 1799 AGTGCTGCTCTCAATCTCAATCTCAATCTCGTCTCGTGAGAGAAGTCTCTCGAG 1858
Db 580 AspAlaGlnProLysLys----- 585
QY 1859 AAAGAGGATCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1918
Db 586 -----LysGluGluGlyAsnLysGlyArgLysGlyProLeuSerSerIleLeuArg 602
QY 1919 GCTTTT 1924
Db 603 AlaPhe 604

RESULT 14

GLCA, SOYBN

ID GLCA SOYBN STANDARD; PRT; 605 AA.

AC P13916;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE Beta-conglycinin, alpha chain precursor.

OS Glycine max (Soybean).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

OC Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

OC Glycine.

OX NCBI_TaxID=3847;

RN {1}

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Cotyledon;

RX MEDLINE=91355860; PubMed=2103438;

RA Sebastiani F.L., Farrel L.B., Schuler M.A., Beachy R.N.;

RT "Complete sequence of a CDNA of alpha subunit of soybean beta-

RT conglycinin, ";
 RL Plant Mol. Biol. 15:197-201(1990).
 CC -!- FUNCTION: Seed storage protein. Accumulates during seed
 CC development and is hydrolyzed after germination to provide a
 CC carbon and nitrogen source for the developing seedling.
 CC -!- SUBUNIT: The alpha'-, alpha, and beta-subunits associate in
 CC various combinations to form trimeric proteins.
 CC -!- SUBCELLULAR LOCATION: Embryo axis, and cotyledonary membrane-bound
 CC vacuolar protein bodies.
 CC -!- SIMILARITY: Belongs to the 7S seed storage protein family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL; X17698; CAA35691.1; -; mRNA.
 DR PIR; S14681; FWSYBA.
 DR HSSP; P25974; IIPJ.
 DR SMK; P13916; 194-581.
 DR InterPro; IPR006045; Cupin.
 DR InterPro; IPR007113; Cupin region.
 DR Pfam; PF00190; Cupin_1; 2.
 KW Glycoprotein; Multigene family; Seed storage protein; Signal;
 KW Storage protein; Vacuole.
 FT SIGNAL 1 22
 FT PROPEP 23 62
 FT CHAIN 63 605 Beta-conglycinin, alpha chain.
 FT CARBOHYD 261 261 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 517 517 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 605 AA; 70293 MW; CBEBAS0506BBBC57 CRC64;
 Alignment Scores:
 Pred. No.: 5,22e-76 Length: 605
 Score: 1281.00 Matches: 285
 Percent Similarity: 59.2% Conservative: 107
 Best Local Similarity: 43.1% Mismatches: 174
 Query Match: 35.7% Indels: 96
 DB: 1 Gaps: 18
 US-10-728-323-1 (1-2032) x GLCA_SOYBN (1-605)
 QY 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTTCTAGGATCTTGTCTGGGTTTCAGTT 109
 DB 2 MetArgAlaArgPhePro-----LeuLeuLeuGluGlyLeuValPheLeuAlaSerVal 19
 QY 110 TCTGCAACGCATGCCAAGTCATCACCTTACCAGAAGAAACAGAGAACCCCTCGCCCCAG 169
 DB 20 SerValSerPheGlyIleAla-----TyrTrpGluLys-----GluAsnProLysHisAsn 36
 QY 170 AGTGGCTCCAGAGTTGTCAACAGGACCGGATGACTTGAAGCAAAAGGCATCGGATCT 229
 DB 37 LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 56
 QY 230 CGCTGCACCAAGCTCAGATGAT-----CTCGTTGT 262
 DB 57 ArgCysAsnLeuLeuLysValGluLysGluGluCysGluGluGlyLeuIleProArg--- 75
 QY 263 GTCTATGATCCTCGAGGACACTGGCACCACCAACCAAGCTTCCCTCCAGGGGAGCGG 322
 DB 76 -----ProArgProArg-----ProGlnHisProGlu 84
 QY 323 ACACGTGGCGCAACCCGAGACTACGATGATGAC----- 358
 DB 85 ArgGluProGlnProGlyGluLysGluLysGluLysGluLysGluLysGluLysGluLys 104
 QY 359 -----CGCGTCNACCCCAAGAGAG-----GNAGAGGCCGA 391
 DB 105 ProPheProArgProGlnProArgGlnGlnGluHisGluGlnArgGluGluGlnGlu 124
 QY 392 TGG-----GGACCAAGCTGGACCGAGGGAGCGTGAAGAGAGAA 430
 DB 471 AsnIleGluLeuValGlyLeu---LysGluGlnGlnGlnGlnGlnGlnGlnGlnGln 489

Db 125 TrpProArgLysGluGluLysArgGlyGluLysGlySerGluGluGluAspGluAspGlu 144
 QY 431 GACTGAGAGACAACCAAGAGAGAT-----TGGAGCGCACCAAGTCATCATCAGCAGCCACGG 484
 Db 145 AspGluGluGlnAspGluArgGlnPheProPheProArgProHisGlnLysGluGlu 164
 QY 485 AAAATAAGGCCCGCAAGGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 544
 Db 165 ArgAsnGluGluGluAspGluGluGlnArgGluGluGluGluGluGluGluAsp 184
 QY 545 AGGGAA-----GAAACATCTCGGAACACCCCTTTCTACTCCCGTCAAGCGCGTTTACG 598
 Db 185 SerGluLeuArgArgHisLysAsnLysAsnProPheLeuPheGlySerAsnArgPheGlu 204
 QY 599 ACCCGCTACGGGAACCAAAACCGTAGGATCCGGTCTCGCAGAGGTTTCCACCAAGGTC 658
 Db 205 ThrLeuPheLysAsnGlnTyrGlyArgIleArgValLeuGlnArgPheAsnGlnArgSer 224
 QY 659 AGGCAGTTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTAACACT 718
 Db 225 ProGlnLeuGlnAsnLeuArgAspTyrArgIleLeuGluPheAsnSerLysProAsnThr 244
 QY 719 CTGTGTTCTTCCCAAGCACCGCTGATGCTGATAACATCTTTGTTATCCAGCAAGGCGCAAGCC 778
 Db 245 LeuLeuLeuProAsnHisAlaAspAlaAspTyrLeuIleValIleLeuAsnGlyThrAla 264
 QY 779 ACCGTGACCGTAGCAATGGCAATACAGAAAGAGCTTTAATCTTCACGAGGCGCCATGCA 838
 Db 265 IleLeuSerLeuValAsnAsnAspArgAspSerTyrArgLeuGlnSerGlyAspAla 284
 QY 839 CTCAGAAATCCCATCCGGTTTCATTTCTCTACATCTTGAACCGCATCACACCAAGAACCTC 898
 Db 285 LeuArgValProSerGlyThrTyrTyrValValAsnProAspAsnAsnGluAsnLeu 304
 QY 899 AGAGTAGCTAAAATCTCCATGCCGTTTAAACACACCCCGCCAGTTTGAGGATTTCTTCCCG 958
 Db 305 ArgLeuIleThrLeuAlaIleProValAsnLysProGlyArgPheGluSerPhePheLeu 324
 QY 959 GCGAGCAGCGGAGACCAATCATCTTCTGAGGGCTTCAGAGGATAGCTTGGAGGCC 1018
 Db 325 SerSerThrGluAlaGlnGlnSerTyrLeuGlnGlyPheSerArgAsnIleLeuGluAla 344
 QY 1019 GCCTTCAATCGGAATTCAATGAGATACGAGGGTCTCTTAGAAGAGAATCGCAGGAGGT 1078
 Db 345 SerTyrAspThrLysPheGluIleAsnLysValLeuPheSerArgGluGly--- 363
 QY 1079 GAGCAAGAGAGAGAGGCGAGCGAGTACTCGGAGTAGTGAGAACATGAAGGA 1138
 Db 364 -----GlnGlnGlnGlyGluGlnArgLeu-----GlnGluSer 374
 QY 1139 GTGATAGTCAAGTGTCAAAGGAGACGTTTGAAGAACTTACTAAGCACGCTAAATCCGTC 1198
 Db 375 ValIleValIleSerLysGluGlnIleArgAlaLeuSerLysArgAlaLysSerSer 394
 QY 1199 TCAAGAAAAGGCTCCCAAGAGAGGAGAGATATCACCAACCCCACTCACTTGAGAGAGGC 1258
 Db 395 SerArgLysThrIleSerSerGlu-----AspLysProPheAsnLeuArgSerArg 411
 QY 1259 GAGCCCGATCTTTCTAAACAACCTTTGGGAAGATTATTGAGGTGAAGCAGACAAAGAGAAC 1318
 Db 412 AspProIleTyrSerAsnLysLeuGlyLysPhePheGluIleThrProGlu---LysAsn 430
 QY 1319 CCCAGCTTCAGACCTGGACATGATGCTCACCTGTGTAGATCAAGAGAGAGCTTTC 1378
 Db 431 ProGlnLeuArgAspLeuAspIlePheLeuSerIleValAspMetAsnGluGlyAlaLeu 450
 QY 1379 ATGCTCCCACTTCAACTCAAAGGCGCATGTTTATCGTCTCGTCTCAACAAAGAACTGGA 1438
 Db 451 LeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGlyAspAla 470
 QY 1439 AACCTTGAATCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGACGCGGGGAAGAGAG 1498
 Db 471 AsnIleGluLeuValGlyLeu---LysGluGlnGlnGlnGlnGlnGlnGlnGlnGln 489

QY 1499 GAGGACCAAGACGAAGAAGAGGAGGAAGTAACAGAGAGGTGGTAGGTACACAGCGAGG 1558
 Db 490 Pro-----LeuGluValArgLysTyrArgAlaGlu 499
 QY 1559 TTGAAGGAAGGCGATGTGTATCATCGCAGCAGCTCATCCAGTAGCCATCAACGCTTCC 1618
 Db 500 LeuserGluGlnAspIlePheValIleProAlaGlyTyrProValValValAsnAlaThr 519
 QY 1619 TCGGAATCCATCTGCTGGCTTCGATCAAGCTGAAACACACACAGAAATCTTCCTT 1678
 Db 520 SerAsnLeuAsnPhePheAlaIleGlyIleAsnAlaGluAsnGlnArgAsnPheLeu 539
 QY 1679 GCAGGTGATAAGGACAAATGTGATAGACCAAGATAGAGAAGCAAGCGAAGGATTAGCATTC 1738
 Db 540 AlaGlySerGlnAspAsnValIleSerGlnIleProSerGlnValGlnGluLeuAlaPhe 559
 QY 1739 CTGGGTGGGTGAACAAGTTGAGAAGCTCATCAAAAACAGAGAATCTCACTTTGTG 1798
 Db 560 ProGlySerAlaGlnAlaValGluLysLeuLeuLysAsnGlnArgLysSerTyrPheVal 579
 QY 1799 AGTGCTCGTCCCTCAATCTCAATCTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTGAG 1858
 Db 580 AspAlaGlnProLysLys-----585
 QY 1859 AAGAGGATCAAGAGGAGGAAACCAAGGAGGAGGTCCACTCTTTCAATTTTGAAG 1918
 Db 586 -----LysGluGluGlyAsnLysGlyArgLysGlyProLeuSerSerIleLeuArg 602
 QY 1919 GCTTTT 1924
 Db 603 AlaPhe 604
 RESULT 15
 CVCA_PEA STANDARD; PRT; 571 AA.
 AC P13915;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Convicilin precursor.
 GN Names=CVCA;
 OS Pisum sativum (Garden pea).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC rosids; eurosids I; fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
 OX NCBI_TaxID=3888;
 RN [1]
 RP NUCLEOTIDE SEQUENCE, AND PARTIAL PROTEIN SEQUENCE.
 RC STRAIN=cv. Feltham First;
 RX MEDLINE=88326208; PubMed=3415641;
 RA Bown D., Ellis T.H.N., Gatehouse J.A.;
 RT "The sequence of a gene encoding convicilin from pea (Pisum sativum
 L.) shows that convicilin differs from vicilin by an insertion near
 the N-terminus."
 RL Biochem. J. 251:717-726(1988).
 CC -!- FUNCTION: Seed storage protein.
 CC -!- SUBCELLULAR LOCATION: Cytosolic membrane-bound vacuolar protein
 CC bodies.
 CC -!- SIMILARITY: Belongs to the 7S seed storage protein family.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL; X06398; CAA29695.1; -; Genomic_DNA.
 DR PIR; S00566; S00566.
 DR HSSP; P25974; IIRU.
 DR InterPro; IPR006045; Cupin.
 DR InterPro; IPR007113; Cupin_region.
 DR Pfam; PF00190; Cupin_1; 2.

KW Direct protein sequencing; Multigene family; Seed storage protein;
 KW SIGNAL; Storage protein; Vacuole.
 FT SIGNAL 1 28
 FT CHAIN 29 571 Convicilin.
 SQ SEQUENCE 571 AA; 66990 MW; 749CFBEB2D16D57B CRC64;

Alignment Scores:
 Pred. No.: 4,34e-75 Length: 571
 Score: 1267.00 Matches: 276
 Percent Similarity: 56.7% Conservative: 98
 Best Local Similarity: 41.8% Mismatches: 146
 Query Match: 35.3% Indels: 140
 DB: 1 Gaps: 13

US-10-728-323-1 (1-2032) x CVCA_PEA (1-571)

QY 68 CCACGTGATGCTTCTCTAGGATCTTGTCTCTGGCTTCAAGTCTGCAACGATGCCAAG 127
 Db 10 ProLeuLeuPheLeuGlyIlePheLeuAlaSerValCysValThrTyrAlaAsn 29
 QY 128 TCATCATCTTACCAGAGAAACAGAAACCCCTGCGCCAGAGGTGCTCCAGAGTTGT 187
 Db 29 -----29
 QY 188 CAACAGGAACCGGATGACTTGAAGCAAAAGGCATCGAGTCTCGTGCACCAAGTCGAG 247
 Db 29 -----29
 QY 248 TATGATCTCTGTTGTCTATGATCTCTGAGGACACACTGGCCACCACCAACACGTTCC 307
 Db 30 -----TyrAspGluGlySerGluThrArgVal-----38
 QY 308 CCTCAGGAGCGGACACACTGGCCGCAACCCGAGACTACGATGATGACCGCGTCAA 367
 Db 39 -----ProGlyGlnArgGluArgGlyArgGlnGluGlyGluGluLysArgHis---56
 QY 368 CCCCAGAGAGAGAGGAGCGGATGGGACACGACTGGACCGAGGAGCGT-----418
 Db 57 -----GlyGluTrpArgProSerTyrGluLysGluGluHisGluGlu 70
 QY 418 -----418
 Db 71 GluLysGlnLysTyrArgTyrGlnArgGluLysLysGluLysGluValGlnProGly 90
 QY 419 -----GAAAGAGAAGAACTGGAGACACCAAGAGAGATTGGAGG-----460
 Db 91 ArgGluArgTrpGluArgGluAspGluGlnValGluGluLysArgGlySer 110
 QY 461 -----CGACCAAGTCATCAGCAGCCACCG 484
 Db 111 GlnArgArgGluAspProGluGluArgAlaArgLeuArgHisArgGluArgThrLys 130
 QY 485 AAAATAAGGCCGAGGAG 544
 Db 131 ArgAspArgArgHisGlnArgGluGluGluGluGluGluGluGluGluGluGluGlu 145
 QY 545 AGGGAAGAAACATCTCGGAACCAACCTTCTACTCTCCCTCAAGGGCGTTTACACCCGC 604
 Db 146 SerGluSerGlnGluHisArgAsnProPheLeuPheLysSerAsnLysPheLeuThrLeu 165
 QY 605 TAGCGGAACCAACCGTAGGATCCGGGTCTTCGAGAGGTTTGACCAAAAGTCAAGCGCAG 664
 Db 166 PheGluAsnGluAsnGlyHisIleArgLeuGlnArgPheAspLysArgSerAspLeu 185
 QY 665 TTTCAGAAATCTCAGAAATCACCGTATTGTGAGATCGAGGCCCAACCTAACACTCTTGT 724
 Db 186 PheGluAsnLeuGlnAsnTyrArgLeuValGluTyrArgAlaLysProHisThrIlePhe 205
 QY 725 CTTCCCAAGCAGCTGATGCTGATTAACATCTTGTATCCAGCAAGGCGAAGCCACCGTG 784
 Db 206 LeuProGlnHisIleAspAlaAspLeuIleLeuValLeuAsnGlyLysAlaIleLeu 225
 QY 785 ACCGTAGCAATGCGCAATAACAGAAAGAGCTTTAATCTTTGACGAGGGCCATGCACCTCAGA 844

```
Db      226 ThrValLeuSerProAsnAspArgAsnSerTyrAsnLeuGluArgGlyAspThrIleLys 245
      ||||| ::: ||:::||| ||:::|||||::: ||| ::::
QY      845 ATCCCATCCGGTTTCATTTCTCATCTTGAACCGCCATGACACACAGAACCTCAGAGTA 904
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      246 IleProAlaGlyThrSerTyrLeuValAsnGlnAspAspGluGluAspLeuArgVal 265
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY      905 GCTAAATCTCCATGCCGTTAAACACACCCGGCCAGTTTCAGGATTTCTCCCGGCGAGC 964
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      266 ValAspPheValIleProValAsnArgProGlyLysPheGluAlaPhe-----GlyLeu 283
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY      965 AGCCGAGCAACATCATCTTCTGAGGGCTTCAGCAGGAATACGTTGGAGGCGCGCTTC 1024
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      284 SerGluAsnLysAsnGlnTyrLeuArgGlyPheSerIysAsnIleLeuGluAlaSerLeu 303
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY      1025 AATCGGAATTCATGAGATACGAGGGTCTCTTAGAAGAGAAATCGAGAGGTGAGCAA 1084
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      304 AsnThrLysTyrGluThrIleGluLysValLeuLeuGluGlnGluLysProGln 323
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY      1085 GAGGAGAGGGCAGAGCGCATGGAGTACTCGGAGTAGTGAGAACATGAAGAGTGATA 1144
      :||:||||:||||:||||:||||:||||:||||:||||:||||:||||
Db      324 GlnLeuArgAspArgLysArg-----ThrGlnGlnGlyGluGluArgAsp---AlaIle 340
      :||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY      1145 GTCAAAGCTGTCAAAGGAGCAGCTTGAAGAACTTACTAAGCACGCTAAATCCGCTCTCAAAG 1204
      :||:||||:||||:||||:||||:||||:||||:||||:||||:||||
Db      341 IleLysValSerArgGluGlnIleGluGluLeuArgLysLeuAlaLysSerSerLys 360
      :||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY      1205 AAAGGCTCCGAAGAAGGAGGAGATATCACCAACCCAATCACTTGAGAGAACGCGAGCCC 1264
      |||||:||||:||||:||||:||||:||||:||||:||||:||||
Db      361 LysSerLeuProSerGluPheGlu-----ProPheAsnLeuArgSerHisLysPro 377
      |||||:||||:||||:||||:||||:||||:||||:||||:||||
QY      1265 GATCTTTCTAAACACTTTGGGAAGTTATTTGAGGTGAAGCCAGACACAAGAAACCCCCAG 1324
      :||:||||:||||:||||:||||:||||:||||:||||:||||:||||
Db      378 GluTyrSerAsnLysPheGlyLysLeuPheGluIleThrProGluLysLysTyrProGln 397
      :||:||||:||||:||||:||||:||||:||||:||||:||||:||||
QY      1325 CTTCAGGACCTGACATGATGCACCTGTGTAGAGATCAAGAAGGAGCTTTGATGCTC 1384
      |||||:||||:||||:||||:||||:||||:||||:||||:||||
Db      398 LeuGlnAspLeuAspIleLeuValSerCysValGluIleAsnLysGlyAlaLeuMetLeu 417
      |||||:||||:||||:||||:||||:||||:||||:||||:||||
QY      1385 CCACACTTCAACTCAAAGGCCATGGTTATCGTCGTCTCAACAAAGAACTGGAAACCTT 1444
      |||||:||||:||||:||||:||||:||||:||||:||||:||||
Db      418 ProHisTyrAsnSerArgAlaIleValValLeuLeuValAsnGluGlyLysGlyAsnLeu 437
      |||||:||||:||||:||||:||||:||||:||||:||||:||||
QY      1445 GAATCTCGTCTGTAGAAAAGACACACAGAGGGGACGGCGGGNAGAGGAGGAC 1504
      |||||:||||:||||:||||:||||:||||:||||:||||:||||
Db      438 GluLeuLeuGlyLeuLysAsnGluGlnGlnGluArg-----449
      |||||:||||:||||:||||:||||:||||:||||:||||:||||
QY      1505 GAAGACGAAGAAGAGGGAAGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTTGAAG 1564
      |||||:||||:||||:||||:||||:||||:||||:||||:||||
Db      450 GluAspArgLysGluArg-----AsnAsnGluValGlnArgTyrGluAlaArgLeuSer 467
      |||||:||||:||||:||||:||||:||||:||||:||||:||||
QY      1565 GAAGGCGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGTTCTCCGAA 1624
      |||||:||||:||||:||||:||||:||||:||||:||||:||||
Db      468 ProGlyAspValValIleIleProAlaGlyHisProValAlaIleSerAlaSerSerAsn 487
      |||||:||||:||||:||||:||||:||||:||||:||||:||||
QY      1625 CTCATCTGTTGGCTTCGGTATCAACGCTGAACAAACACACAGAACTCTTCCTGCAGGT 1684
      |||||:||||:||||:||||:||||:||||:||||:||||:||||
Db      488 LeuAsnLeuLeuGlyPheGlyIleAsnAlaLysAsnAsnGlnArgAsnPheLeuSerGly 507
      |||||:||||:||||:||||:||||:||||:||||:||||:||||
QY      1685 GATAAGGACAAATGTGATAGACCATAGAGAAGCAAGCAAGGATTTTAGCATTTCCCTGGG 1744
      |||||:||||:||||:||||:||||:||||:||||:||||:||||
Db      508 SerAspAspAsnValIleSerGlnIleGluAsnProValLysGluLeuThrPheProGly 527
      |||||:||||:||||:||||:||||:||||:||||:||||:||||
QY      1745 TCGGGTGAACAAGTTGAGAGCTCATCAAAAACAGAAAGGAATCTCACTTTGTGAGTGCT 1804
      |||||:||||:||||:||||:||||:||||:||||:||||:||||
Db      528 SerSerGlnGluValAsnArgLeuIleLysAsnGlnLysGlnSerHisPheAlaSerAla 547
      |||||:||||:||||:||||:||||:||||:||||:||||:||||
QY      1805 CGTCTCAATCTCAATCTCAATCTCCGTCGTCCTCGAAGAGAGTCTCCTGAGAAAGAG 1864
      |||||:||||:||||:||||:||||:||||:||||:||||:||||
Db      548 GluPro-----Glu 550
      |||||:||||:||||:||||:||||:||||:||||:||||:||||
QY      1865 GATCAAGAGGAGGAAAAACCAAGGAGGAGGTCCTCTTTCATTTTGAAGGCTTTT 1924
      :||:||||:||||:||||:||||:||||:||||:||||:||||:||||
```

Db 551 GlnLysGluGluSerGlnArgLysArgSerProLeuSerSerValLeuAspSerPhe 570

Search completed: May 15, 2006, 22:18:16
Job time : 502.942 secs

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 23:34:07 ; Search time 57.3506 Seconds
(without alignments)
4441.255 Million cell updates/sec

Title: US-10-728-323-1

Perfect score: 3586

Sequence: 1 aataatcatatatattcatc.....cgttgtgcgtttttctcc 2032

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p model -DRV=xlp
-Q=/abs/ABSWEB_spool/US10728323/runat_15052006_172145_22607/app_query.fasta_1
-DB=Published Applications AA_Main -QFMT=fastan -SUFFIX=n2p.rapbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss02p
-USER=US10728323 @CGN_1_1_503 @runat_15052006_172145_22607 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA_Main.*

- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pcp.*
- 2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pcp.*
- 3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pcp.*
- 4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pcp.*
- 5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pcp.*
- 6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3286	91.6	626	3	US-09-847-208-28
2	3286	91.6	626	4	US-10-228-806-2
3	3286	91.6	626	4	US-10-100-303A-7
4	3286	91.6	626	4	US-10-245-871-10
5	3286	91.6	626	4	US-10-253-286-10
6	3286	91.6	626	5	US-10-809-689-95
7	3286	91.6	626	5	US-10-899-551-2
8	3192	89.0	635	5	US-10-899-551-53
9	3052	85.1	634	3	US-09-731-221-78
10	3041	84.8	614	3	US-09-331-631A-21
11	3041	84.8	614	3	US-09-847-208-27

ALIGNMENTS

RESULT 1

US-09-847-208-28
; Sequence 28, Application US/09847208
; Publication NO. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; FILE REFERENCE: IGE-MEDIATED ALLERGIC DISEASES
; FILE REFERENCE: UC67.002A
; CURRENT APPLICATION NUMBER: US/09/847,208
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Arachis hypogaea (Peanut)
US-09-847-208-28

Alignment Scores:
Pred. No.: 2.56e-261 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.6% Indels: 0
DB: 3 Gaps: 0

US-10-728-323-1 (1-2032) x US-09-847-208-28 (1-626)

QY 50 ATGAGAGGAGGTTTCTCCACTGATGCTGTGTAGGATCCTTGTCTGCTTCAGTT 109

Db 1 MetArgGlyArgValserProLeuMetLeuGlyLeuValLeuAlaSerVal 20

Sequence 21, Appli
Sequence 8, Appli
Sequence 55, Appli
Sequence 153195,
Sequence 25, Appli
Sequence 110, Appl
Sequence 8840, Ap
Sequence 153206,
Sequence 260105,
Sequence 14, Appli
Sequence 12, Appli
Sequence 2, Appli
Sequence 4408, A
Sequence 51703, A
Sequence 3, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 43905, A
Sequence 49262, A
Sequence 43971, A
Sequence 49373, A
Sequence 8, Appli
Sequence 43597, A
Sequence 43984, A
Sequence 51411, A
Sequence 54, Appli
Sequence 7, Appli
Sequence 7, Appli

614 84.8 3041 614 4 US-10-147-095-21
614 84.8 3041 614 4 US-10-100-303A-8
268 38.9 1394 268 4 US-10-100-303A-55
605 35.8 1282 605 4 US-10-424-599-153195
605 35.7 1281 605 3 US-09-331-631A-25
605 35.7 1281 605 4 US-10-147-095-25
623 35.6 1275.5 623 5 US-10-424-599-153206
623 35.5 1274.5 623 5 US-10-739-930-8840
605 35.4 1268 605 4 US-10-100-303A-110
584 35.3 1201.5 584 4 US-10-424-599-260105
417 32.3 1157 417 4 US-10-245-227B-14
425 32.3 1157 425 4 US-10-245-227B-12
439 32.3 1156 439 4 US-10-245-227B-1
390 31.2 1117.5 390 4 US-10-245-227B-2
324 24.4 874 324 4 US-10-425-114-44408
344 24.2 867 344 4 US-10-425-114-51703
666 24.1 865.5 666 3 US-09-331-631A-3
666 24.1 865.5 666 4 US-10-147-095-3
625 24.1 863.5 625 3 US-09-331-631A-5
625 24.1 863.5 625 4 US-10-147-095-5
666 23.7 849.5 666 3 US-09-331-631A-1
666 23.7 849.5 666 4 US-10-147-095-1
316 23.1 828 316 4 US-10-425-114-43905
316 23.1 828 316 4 US-10-425-114-49262
315 23.1 827 315 4 US-10-425-114-43971
301 22.8 816 301 4 US-10-425-114-49373
590 22.1 791 590 3 US-09-331-631A-8
590 22.1 791 590 4 US-10-147-095-8
300 22.0 790 300 4 US-10-425-114-43597
296 22.0 788 296 4 US-10-425-114-43984
291 21.5 770 291 4 US-10-425-114-51411
149 21.2 761 149 4 US-10-100-303A-54
525 20.6 737.5 525 3 US-09-331-631A-7
525 20.6 737.5 525 4 US-10-147-095-7

QY 110 TCTGCAACGATCCCAAGTCATCACCTTACCAGAGAAACACAGAGAACCCCTCGCCGACG 169
Db 21 SerAlaThrHisAlaLysSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATCGAGTCT 229
Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGTGCACCAAGCTCGAGTATGATCTCTCGTTGTCTATGATCCTCGAGGACACACTGGC 289
Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAAGTTCCCTCCAGGGGACGACACGTGGCCGCCCAACCCGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCTCAACCCCAAGAGAGGAAGGAGCGGATGGGACCCAGCTGGACCG 409
Db 101 AspAspArgArgGlnProArgArgGluGlyArgGlyProAlaGlyPro 120
QY 410 AGGAGCGTGAAGAGAAGACTGGAGACAAACCAAGAGAAGATTGGAGGCGACCAAGT 469
Db 121 ArgGluArgGluArgGluGluAspTyrArgGlnProArgGluAspTyrArgProSer 140
QY 470 CATCAGCAGCCAGGAAATAAGCCCGAAGGAGAGAGAGAGAACCAAGAGTGGGAACA 529
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTyrGlyThr 160
QY 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAAACCTTCTACTTCCCGTCAAGG 599
Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTACGACCCGCTACGGGAACCAAAACGGTAGGATCCGGTCTCTGCAGAGGTTTGAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGTCAAGGCAGTTTCAGATCTCCAGAAATCCACGTTATGTGCAGATCGAGGCCAAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTGTCTCCCAAGCAGCGCTGATGCTGATAACATCTTGTATTCCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTGACCGTAGCAAAATGGCAATTAACAGAAAGACTTAACTGTGACGAG 829
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCATGCACTCAGNATCCCATCCGTTTTCATTTCTTACATCTTGACCCGCTGACAAC 889
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCCGTTTAAACACACCCGCCAGTTTGAGGAT 949
Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCGAGCAGCGAGACCAATCATCTTACTTTCAGGCTTTCAGCAGATAACG 1009
Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGCCGCTTCATGCGGAATTCATAGATAGCGAGGCTGCTGTAGAGAGAAT 1069
Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGGTGAGCAAGAGGAGAGGCGCAGGCGATGGAGTACTCGGAGTAGTGAGAAC 1129
Db 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTyrPheSerSerSerGluAsn 360
QY 1130 AATGAAGAGTATAGTCAAGTGTCAAAGTCAAAGAGCACGTTGAAGAACTTACTAAGACGCT 1189
Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AAATCCGCTCTCAAAGAAAGGCTCCGAGAGAGAGGAGATATCACCACCAACCAATCACTTG 1249

Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAGAAGGCGAGCCCGCATCTTCTTAACAACATTTGGGAAGTTATTTGAGGTGAAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AGAGAAGAACCCCGAGCTTCAGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAAGGCATGTTATCTGCTCGTCAACAAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
QY 1430 GGAACCTGGAACCTTGAACCTCGTGGCTGTGAAGAAAGACAAACACAGAGGGGACCGCG 1489
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgTyr 480
QY 1490 GAAGAAGAGGAGACGAAACGAAAGAGAGGAGGAGTAACACAGAGAGGTGCGTAGGTAC 1549
Db 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTTGAAGGAAGCGATGTGTTTCATCATGCCACGAGCTCATCCAGTAGGCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGAACTCCATCTGCTTGGCTTTCGCTATCAACGCTCAAAACACACACAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTCGAGGTGATTAAGACAAATGTATAGACCATAGACAGGAAAGCGAAGGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCAATTCCTCGGTGCGGTGAACAAAGTTGAGAAGCTCATCAAAACACAGAGGAATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTCTCGTCTCTCAATCTCAATCTCAATCTCCGTCGCTCTCTGAGAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTCAGAAAGAGATCAAGAGGAGAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
Db 601 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTCGAAGCTTTTAAC 1927
Db 621 IleLeuLysAlaPheAsn 626

RESULT 2

US-10-228-806-2
; Sequence 2, Application US/10228806
; Publication No. US20030049237A1
; GENERAL INFORMATION:
; APPLICANT: Bannon, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reactions
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0043
; CURRENT APPLICATION NUMBER: US/10/228, 806
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-228-806-2

Alignment Scores: 2,56e-261 Length: 626
Pred. No.: 3286.00 Matches: 626
Score:

Percent Similarity: 100.0%		Conservative: 0	
Best Local Similarity: 100.0%		Mismatches: 0	
Query Match: 91.6%		Indels: 0	
DB: 4		Gaps: 0	
US-10-728-323-1 (1-2032) x US-10-228-806-2 (1-626)			
Qy	50	ATGAGAGGAGGGTTTCTCCACTGATGCTGTTCTAGGATCCTTGCTCGCTTCAGTT	109
Db	1	MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal	20
Qy	110	TCTGCAACGATGCCAAGTCATCACCTTACCAGAGAAACAGAGAACCCCTGCGCCAG	169
Db	21	SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln	40
Qy	170	AGTGCCCTCCAGAGTTGTCAACAGGAACCGGATGCTTGAAGCAAAAGGCATGCGAGTCT	229
Db	41	ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer	60
Qy	230	CGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGGC	289
Db	61	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly	80
Qy	290	ACCACCAACCAAGTTCCCTCCAGGGAGCGGACACGTGCGCCGCCAACCCGGAGACTAC	349
Db	81	ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	100
Qy	350	GATGATCACCGCTCAACCCCGAAGAGAGGAGGAGCGGATGGGAGCCAGCTGGACCG	409
Db	101	AspAspArgArgGlnProArgGluGluGlyArgGluGluGlyArgGluGluGluPro	120
Qy	410	AGGAGCGTGAAGAGAAAGAACTGGAGACCAACAGAGAGATTGGAGGCGACCAAGT	469
Db	121	ArgGluArgGluArgGluGluAspTyrArgGlnProArgGluAspTyrArgArgProSer	140
Qy	470	CATCAGCAGCCCGAAATAAGCCCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	529
Db	141	HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr	160
Qy	530	CCAGGTAGCCATGTGAGGAGAAACATCTCGGAACAACCTTCTACTTCCCGTCAAGG	589
Db	161	ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg	180
Qy	590	CGGTTTAGCACCCGCTACGGGAACCAAAACGGTAGGATCCGGGCTCTCGAGAGTTTGAC	649
Db	181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	200
Qy	650	CAAGGTCAAGCAGTTTCCAGATCTCCAGATCACCGTATTGTGCAGATCCGAGGCCAAA	709
Db	201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys	220
Qy	710	CCTAACACTCTTGTCTTCCCAAGCAGCTGATGCTGATAACATCTTGTATCCAGCAA	769
Db	221	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	240
Qy	770	GGCAAGCCACCGTAGCAGCAATGCGCAATTAACAGAAAGAGCTTTAATCTTACGAG	829
Db	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260
Qy	830	GGCATGCATCAGAAATCCCATCCGGTTTCATTTCTACATCTTTGAACCCGCTATGCAAC	889
Db	261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	280
Qy	890	CAGAACCTCAGAGTAGTAAATCTCCATGCCGCTTAAACACACCCGCCAGTTTGAGAT	949
Db	281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	300
Qy	950	TTCTTCCCGGAGCAGCCGAGACCAATCATCTTCTTTCAGGGCTTTCAGAGGAATACG	1009
Db	301	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	320
Qy	1010	TTGAGGCGCCCTTCAATGCGGAATTCATATGATACGAGGGTGTCTGTAGAAGAGAAT	1069

DB	321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn	340
QY	1070	GCAGGAGGTGAGCAAG	1129
DB	341	AlaGlyGlyGluGlnGluArgGlyGlnArgArgGlyGlnArgArgSerGluAsn	360
QY	1130	AATGAAGAGGTATAGTCAAAGTGTCAAGAGAGACAGTGTGAAGAACTTACTAAGCAC	1189
DB	361	AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	380
QY	1190	AAATCCGCTCAAAGAAAGGCTCCCAAGAGAGAGAGAGATATCACCAACCAATCAACTG	1249
DB	381	LysSerValSerLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu	400
QY	1250	ACAGAAAGGCGAGCCCGATCTTCTTCAACAACTTTGGAACTTATTTGAGGTGAAGCCAG	1309
DB	401	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	420
QY	1310	AAGAAGAAACCCCGCTTTCAGGACCTGACATGATGCTCACCTGTGTAGAGATCAAGAA	1369
DB	421	LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu	440
QY	1370	GGAGCTTTGATGCTCCACACTTCAACTCAAAGGCGATGTTATCGTCTGCTCAACAA	1429
DB	441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	460
QY	1430	GGAACTGGAAACCTTGAACCTGCTGTAAAGAAAGAGCAACAACAGAGGGGACGCGG	1489
DB	461	GlyThrGlyAsnLeuLeuValAlaValArgLysGluGlnGlnGlnArgGlyArgArg	480
QY	1490	GAAGAGAGGAGGACGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1549
DB	481	GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr	500
QY	1550	ACAGCGAGGTGAAGAGAGGCGATGTTCATCATCCAGCAGCTCATCCAGTAGCCATC	1609
DB	501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520
QY	1610	AACGCTTCTCCGAACTCCATCTGCTTGGCTTCAACGCTGAAACCAACCAACACAGA	1669
DB	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg	540
QY	1670	ATCTTCTTTCAGGTGATAAGGACAATGTGATAGACCAGATAGAGAGAGAGAGAGAT	1729
DB	541	IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp	560
QY	1730	TTAGCATTCCTGGCTCGGTGAAACAGTTGAGAGAGCTCATCAAAAACAGAGAGATCT	1789
DB	561	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer	580
QY	1790	CACCTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCCGCTCTCTCTGAGAAAGAG	1849
DB	581	HisPheValSerAlaArgProGlnSerGlnSerProSerSerProGlyLysGlu	600
QY	1850	TCTCTGAGAAAGAGATCAAGAGAGGAGAAACCAAGAGAGGAGAGGTCCTCTTCA	1909
DB	601	SerProGluLysGluAspGlnGluGluAsnGlnGlyLysGlyProLeuLeuSer	620
QY	1910	ATTTTGAAGGCTTTTAAAC	1927
DB	621	IleLeuLysAlaPheAsn	626

RESULT 3

US-10-100-303A-7
; Sequence 7, Application US/10100303A
; Publication No. US20030202980A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0166
; CURRENT APPLICATION NUMBER: US/10/100,303A
; CURRENT FILING DATE: 2002-03-18

; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Atachis hypogaea, Prot/Nucleo-Ara h 1
US-10-100-303A-7

Alignment Scores:

Pred. No.: 2,566-261 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.6% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-1 (1-2032) x US-10-100-303A-7 (1-626)

QY	50	ATGAGAGGGAGGTTTCTCCACTGATGCTGTTCCTAGGGATCCCTTCTGGCTTCAGTT	109
Db	1	MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal	20
QY	110	TCTGCAACGATGCCAAGTCATCACCTTACCAGAGAAACACAGAGAACCCCTGCGCCACG	169
Db	21	SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln	40
QY	170	AGTGTCTCAGAGTTGTCAACAGGAAACCGGATGACTTGAAGCAAAAGGCGATCGAGTCT	229
Db	41	ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer	60
QY	230	CGTGTCCACCAAGCTCCAGTATGATCTCGTTGTCTATGATCCTCAGGACACACTGGC	289
Db	61	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly	80
QY	290	ACCACCAACCAACGTTCCCTCCAGGGAGCGGACACGTCGGCCCAACCCCGGAGACTAC	349
Db	81	ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	100
QY	350	GATGATGACCCGCGTCAACCCGGAAGAGGAAGGAGCGCGATGGGACCCAGCTGACCG	409
Db	101	AspAspAspArgGlnProArgGluGluGlyArgTrpGlyProAlaGlyPro	120
QY	410	AGGAGCGTCAAGAGAGAGAGACTGGAGACACCAAGAGAACTTGGAGCGGACCAAGT	469
Db	121	ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgProSer	140
QY	470	CATCAGCAGCCCGGAAAATAAGCCCGAAGGAGAGAGAGAACAAAGAGTGGGAAACA	529
Db	141	HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr	160
QY	530	CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAAACCCCTTCTACTTCCCGTCAAGG	589
Db	161	ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg	180
QY	590	CGGTTTAGCACCCGTACCGGAACCAAAACGGTAGGATCCGGTCTCTGCAAGAGTTTGAC	649
Db	181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	200
QY	650	CAAAAGGTCAAGCGATTTCCAGATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAA	709
Db	201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys	220
QY	710	CCTAACACTTGTCTTCCAGCAGCAGCTGATGCTGATAACATCTTGTATTCACGACAA	769
Db	221	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	240
QY	770	GGGCAAGCCACCGTAGCGGTAGCAATGGCAATTAACAGAAAGAGCTTTAATCTTGACGAG	829
Db	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260
QY	830	GGCCATGCACTCAGAAATCCCATCCGTTTCATTTCTTACATCTTGAACCCGCCATGACAAC	889
Db	261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	280

QY	890	CAGAACTTCAGAGTAGCTAAAATCTCCATGCCGTTAAACACACCCGCGCAGTTTGAGGAT	949
Db	281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	300
QY	950	TTCTCCCGCGCAGCAGCGAGACCAATCATCTACTTGCAGGGCTTCAGCAGGAATACG	1009
Db	301	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	320
QY	1010	TTGAGGCGCGCTTCAATCGGAAATCAATGAGATACGAGGGTGTCTGTAGAGAGAAAT	1069
Db	321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn	340
QY	1070	GCAGAGGTGACCAAGAGAGAGAGCGGACGAGCGATCGAGTACGCGAGTAGTGAGAAC	1129
Db	341	AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn	360
QY	1130	AATGAAGAGTAGTATAGTCAAAGTGTCAAAGAGACAGCTTGAAGAACTTACTAAGCACGCT	1189
Db	361	AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	380
QY	1190	AAATCCGTCTCAAGAAAGGCTCCGAAAGAGGAGAGATATCCAAACCCCAATCAACTTG	1249
Db	381	LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu	400
QY	1250	AGAGAGCGCGCCGATCTTCTAACCACTTTGGAGAGTTATTGAGGTGAAGCCAGAC	1309
Db	401	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	420
QY	1310	AAGAAGAACCCCGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA	1369
Db	421	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu	440
QY	1370	GAAGCTTTGATGCTCCACACTTCAACTCAAGAGCCATGGTTATCGTCTGTCAACAAA	1429
Db	441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	460
QY	1430	GGAACTGGAAACCTTCAACTCGTGTAAAGAAAGAGCAACAACAGAGGGGACGCGG	1489
Db	461	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnArgGlyArgArg	480
QY	1490	GAAGAGAGGAGGACGAAGACGAAGAGAGGGAAGTAACAGAGAGGTGCGTAGGTAC	1549
Db	481	GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr	500
QY	1550	ACAGCGAGTTGAGGAAGCGGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC	1609
Db	501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520
QY	1610	AACGCTTCTCCGAACTCCATCTGCTTGGCTTTCGCTATCAACGCTGAAACCAACACAGA	1669
Db	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg	540
QY	1670	ATCTTCTTGCAGGTGATAAGGACAAATGTGATAGACCAAGATAGAGAGCAAGCAAGGAT	1729
Db	541	IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp	560
QY	1730	TTAGCAATTCCTGGGTGCGGTGAAACAAAGTTGAGAGCTCATCAAAAACAGAGGAATCT	1789
Db	561	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer	580
QY	1790	CACTTTGTGAGTGTCTGCTCCTCAATCTCAATCTCCTCAATCTCCGCTGCTCTCTGAGAAGAG	1849
Db	581	HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu	600
QY	1850	TCTCTTGAGAAAGAGATCAAGAGGAGGAGAAAACCAAGGAGGGAAGGCTCCACTCTTTCA	1909
Db	601	SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer	620
QY	1910	ATTTTGAAGGCTTTTAAAC	1927
Db	621	IleLeuLysAlaPheAsn	626

```
RESULT 4
US-10-245-871-10
; Sequence 10, Application US/10245871
; Publication No. US20030235594A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: II-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2013
; CURRENT APPLICATION NUMBER: US/10/245,871
; CURRENT FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-245-871-10

Alignment Scores:
Pred. No.: 2,566-261 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.6% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-1 (1-2032) x US-10-245-871-10 (1-626)
QY 50 ATGAGAGGAGGTTCTCCACTGATGCTGTGCTAGGAGTCTGTCTGCTTCACTT 109
DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGATGCCAAGTCATCACCCTTACCAAGAGAAACAGAGAACCCCTCGCCCGAG 169
DB 21 SerAlaThrHisAlaLysSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAACAAAGCATGCGAGTCT 229
DB 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CCGTCACCAAGCTCGAGTATCATCTCGTGTGCTATGATCTTCGAGGACACACTGGC 289
DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAACGTTCCCTCCAGGGAGCGGACACGTGGCGGCCAACCCGAGACTAC 349
DB 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATACCCCGCTCAACCCCGAAGAGAGAGGAGGCGGATGGGACCAAGCTGGACCG 409
DB 101 AspAspAspArgArgGlnProArgArgGluGluGlyGlyArgTyrGlyProAlaGlyPro 120
QY 410 AGGGACCTGAAGAGAGAAAGACTGGAGACCAACAGAGAGATTTGAGGCGACCAAGT 469
DB 121 ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
QY 470 CATCAGCACCGCAAGAAATAAGGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGlnTrpGlyThr 160
QY 530 CCAGGTAGCCATGTGAGGAGAAACATCTCGGAACAACCCCTTTCTACTTCCCGTCAAGG 589
DB 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCCGCTACGGGAACCAAAACGCTAGATCCGGGTCTTCGAGAGGTTTGC 649
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgGlyArgValLeuGlnArgPheAsp 200
```

```
QY 650 CAAAGGTCAAGCGAGTTCAGAAATCTCCAGAAATCCACCGTATTTGTTCAGATCGAGCCAAA 709
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTGTCTTCCCAAGCAGCTGATGCTGATTAACATCTTGTGTATCCAGCAA 769
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCAGCTGACCGTAGCAATAGCAATAGCAATAGCAATAGCAATAGCAATAGCA 829
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCCATGCACCTCAGAAATCCCATCCGGTTTCATTTCTACATCTTCAACCGCCATGACA 889
DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGAGTAGCTAAAATCTCCATGCCCGTTTAAACACACCCCGCCAGTTGAGN 949
DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCGAGCAGCCGAGACCAATCATCTTCTTTCAGGGCTTTCAGCAGGAGTACG 1009
DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGGCGCGCTTCAATCCGGAATTCAGATACCGAGGGTCTCTTTAGAAGAGAT 1069
DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn 340
QY 1070 GCAGAGGTGAGCAAGAGGAGAGAGGCGAGGCGATGGAGTACTCGGAGTAGTGAGAAC 1129
DB 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgTyrPheSerArgSerGluAsn 360
QY 1130 AATGAAGAGGTAGTCAAAAGTGTCAAAGAGAGACGTTGAAGAACTTACTTAAGCAGCT 1189
DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AAATCCGCTCTCAAAGAAAGGCTCCGAAAGAGGAGATATCACCAACCCCAATCACTTG 1249
DB 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAGAGCGGAGCGCGATCTTTCTAAACACTTTGGGAAGTTATTGTAGGTGAGGCAGAC 1309
DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAAGAACCCCGCTTCCAGACCTGGACATGATCTCCTGTGTAGAGATCAAGAA 1369
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluLysGlu 440
QY 1370 GGAGCTTTGATGCTCCCACTTCAACTCAAAGGCCATGGTTATCTGCTCGTCAACAAA 1429
DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAAGTGAACCTTGAACTCGTGTGCTGTAAAGAAAGCAACAAAGAGGGGACGGCG 1489
DB 461 GlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnGlnArgGlyArgArg 480
QY 1490 GAAGAAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1549
DB 481 GluGluGluGluAspGluAspGluGluGluGluGluGluGluGluGluGluGluGlu 500
QY 1550 ACAGCGAGGTGAAGAGGCGGATGTGTTTCATCATGCCAGCAGCTCATCCAGTACCCATC 1609
DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGAACCTCATCTGCTTGGCTTCGGTATCAACCGCTGAAGAAACCAACAGA 1669
DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTCTGAGGTGATAAGGACAATGTGATAGACCATAGAGAGAGAGAGAGAGAG 1729
DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTTCCCTGGGTGCGGTGAACAGTTTGAGAGGCTCATCAAAACCAAGAGGAATCT 1789
```

Db 561 LeuAlaPheProGlySerGlyGluGlnValGluGlyLeuLeuLeuLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTGCTCGTCTCAATCTCAATCTCAATCTCCGCTCGTCTCTGAGAAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTGTAGAAAGAGATCAAGAGAGAGAAACCAAGGAGGGAAGGGTCCACTCTTTTCA 1909
Db 601 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGCTTTTAAAC 1927
Db 621 IleLeuLysAlaPheAsn 626
RESULT 5
US-10-253-286-10
; Sequence 10, Application US/10253286
; Publication No. US20040058881A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: II-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2015
; CURRENT APPLICATION NUMBER: US/10/253,286
; CURRENT FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 626
; TYPE: PRT
; ORGANISM: *Arachis hypogaea*
US-10-253-286-10
Alignment Scores:
Pred. No.: 2,566-261 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.6% Indels: 0
DB: 4 Gaps: 0
US-10-728-323-1 (1-2032) x US-10-253-286-10 (1-626)
QY 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTTGCTAGGGATCCTTGCTCTGGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGATGCGCAAGTCATCACCTTACAGAGAAACAGAGAACCCCTCGGCCAG 169
Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATCGAGTCT 229
Db 41 ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGTGCACCAAGCTCGAGTATGATCCTCGTGTGTCTATGATCCTCGAGGACACACTGGC 289
Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAACCTTCCCTCCAGGGGAGCGACACGTGCGGCCCAACCCCGGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCCCGCTCAACCCCGAGAGAGGAAGGAGCCGATGGGACACGAGTGGACCG 409
Db 101 AspAspAspArgArgGlnProArgArgGluGluGlyArgTrpGlyProAlaGlyPro 120
QY 410 AGGAGCGCTGAAGAGAGAGAGACTGGAGACCAACAGAGAGATTGGAGCGGACCAAGT 469

Db 121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
QY 470 CATCAGCAGCCACGAAAAATAAGCCCGAAGAGAGAGAGAGAAACAAGAGTGGGAAACA 529
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
QY 530 CCAGGTAGCCATGTGAGGGAGAAACATCTCGGAACAACCCCTTCTACTTCCCGTCAAGG 589
Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCCGCTACGGGAACCAAAACCGGTAGGATCCGGGTCTCTGCAGAGGTTTGAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGTCAAGGCAGTTTCAGAAATCTCCAGAAATCAACCGTATTGTGCAGATCCAGGCCAAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTTGTCTTCCCAAGCAGCGCTGATGCTGATAACATCCTCTTATCCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGCTGACCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAG 829
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCCATGCACTCAGAAATCCCATCCGTTTCATTCTCATCTTGAACCGCCATGACAAC 889
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGAGTAGCTTAAATCTCCATCCCGCTTAAACACACCCGCGCCAGTTGAGGAT 949
Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCGAGCAGCGACCAATCATCTCTTCAGCGGCTTCAGCAGGAATACG 1009
Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGCGCCCTTCAATCGGAATTCATGATGATACGAGGGTGTCTGTAGAGAGAAT 1069
Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGGTGACCAAGAGGAGAGCGGAGCGGATGAGTACTCGGAGTAGTGAGAAC 1129
Db 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGGAGTGATAGTCAAAGTGTCAAAGGACACGTTTCAAAGAACTTACTAAGCACGCT 1189
Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCGTCTCAAAGAAAGGCTCCGAAAGAGGGAGATATACCAACCAATCAACTTG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGNAGAGGCGAGCCCATCTTTCTAACACTTTGGAGTATTATTCAGGTGAAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAAACCCCGAGCTTCAGGACCTCGGACATGATGCTCACCTGTGTAGAGATCAAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATCGGTATTCGTCTCGTCAACAAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAACCTGGAACCTTGAACCTCGGCTGTAAAGAAAGACCAACAAACAGAGGGGACCGCG 1489
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAAGAGAGGAGGACCAAGACCAAGAGAGGAGGAGTAACAGAGAGGTGCGTAGGTAC 1549

Db	401	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLeuPheGluValLysProAsp	420
QY	1310	AGAAGAACCCCGAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA	1369
Db	421	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu	440
QY	1370	GGAGCTTCGATGCTCCACACTTCAACTCAAAAGCCATGGTTATCGTCGTCACAAA	1429
Db	441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	460
QY	1430	GGAACTGGAAACCTTGAACCTCGTGGCTGTAAAGAAAAGAGCAACAACAGAGGGGACGCGG	1489
Db	461	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnArgGlyArgArg	480
QY	1490	GAAGAAGAGAGACGAAAGACGAAGAGAGAGGGAAGTAAACAGAGAGTGCCTAGGTAC	1549
Db	481	GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr	500
QY	1550	ACAGCGAGGTGAAGGAAGCGATGTTCATCATGCCAGCAGCTCATCCAGTAGCCATC	1609
Db	501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520
QY	1610	AACGCTTCCTCCGAACCTCCACTTCGCTTGGCTTCGGTATCAACGCTCAAAAACACACACAGA	1669
Db	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg	540
QY	1670	ATCTTCCTTCAGGTGATAGGACAATGTGATAGACAGATAGAGAACGAAGGAT	1729
Db	541	IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp	560
QY	1730	TTAGCATTCCTCGGTGGTGAACAAGTTGAGAGCTCATCAAAAACAGAGAGGAATCT	1789
Db	561	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer	580
QY	1790	CACCTTGTGAGTGCTCGTCTCAATCTCAATCTCAATCTCCGTCGTCTCCTGAGAAAGAG	1849
Db	581	HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu	600
QY	1850	TTCCTTGAGAAAGAGATCAAGAGGAGGAGAAACCAAGAGGGGAGGTCCTCTTTCA	1909
Db	601	SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer	620
QY	1910	ATTTTGAAGGCTTTTAAAC	1927
Db	621	IleLeuLysAlaPheAsn	626
RESULT 7			
US-10-899-551-2			
; Sequence 2, Application US/10899551			
; Publication No. US20050063994A1			
; GENERAL INFORMATION:			
; APPLICANT: Caplan, Michael J.			
; APPLICANT: Burks, A. Wesley			
; APPLICANT: Sampson, Hugh A.			
; APPLICANT: Howard, Sosin B.			
; APPLICANT: Bottomly, Kim H.			
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy			
; FILE REFERENCE: 2002834-0233			
; CURRENT APPLICATION NUMBER: US/10/899,551			
; CURRENT FILING DATE: 2004-07-26			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 2			
; LENGTH: 626			
; TYPE: PRT			
; ORGANISM: species Atachis hypogea			
US-10-899-551-2			
Alignment Scores:			
Pred. No.:	2,56e-261	Length:	626
Score:	3286.00	Matches:	626
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0

Query Match:	91.6%	Indels:	0
DB:	5	Gaps:	0
US-10-728-323-1 (1-2032) x US-10-899-551-2 (1-626)			
QY	50	ATGAGAGGAGGAGGTTTCTCCACTGATGCTGTGTGCTAGGATCCTTCTCGTGGCTTCAGTT	109
Db	1	MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal	20
QY	110	TGTCAACCATGCCAAGTCATCATCTTACCAGAGAAACACAGAAACCCCTGCGCCACG	169
Db	21	SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln	40
QY	170	AGGTGCTCTCAGAGTTGTCAAAGGNAACCGGTGACTTGAAGCAAAAGCATCGAGTCT	229
Db	41	ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysIleLysAlaCysGluSer	60
QY	230	CGCTGCACCAAGCTCCAGTATGATCCTCGTTGTGTCTATGATCCTCCAGACACACTGGC	289
Db	61	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly	80
QY	290	ACCACCAACCAACGTTTCCCTCCAGGGAGCGGACACGTGGCGCCCAACCCGGAGACTAC	349
Db	81	ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	100
QY	350	GATGATACCGCGCTCAACCCGNAAGAGAGAGGAGCGCGATGGGACCCAGCTGGACCG	409
Db	101	AspAspAspArgArgGlnProArgArgGluGluGlyGlyArgTyrPglyProAlaGlyPro	120
QY	410	AGGAGCGTGAAGAGAAAGAACTCGAGACAAACCAAGAGAAAGATTGGAGGCGACCAAGT	469
Db	121	ArgGluArgGluArgGluGluAspTyrArgGlnProArgLysArgArgProSer	140
QY	470	CATCAGCACCCGCGAAATAAGCCGAGAGAGAGAGAGAGAACCAAGAGTGGGGAACA	529
Db	141	HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTyrGlyThr	160
QY	530	CCAGGTAGCATGTGAGGGAAGAAACATCTCGGAACAACCTTCTTACTTCCCGTCAAGG	599
Db	161	ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg	180
QY	590	CGGTTTGTAGCACCCGCTACCGGAAACCAAAACCGGTAGGATCCGGTCTCGCAGAGTTTGC	649
Db	181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	200
QY	650	CAAAGTCAAGGCAGTTTCAGAAATCTCCAGAAATCACCGTATTTGTGCAGATCGAGGCCAA	709
Db	201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys	220
QY	710	CCTAACACTTGTCTTCTCCAGACGCTGATGCTGATAACATCCTTGTATCCAGCAA	769
Db	221	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	240
QY	770	GGCAAGCCACCGTGACCGTAGCAATGGCAATTAACAGAAAGAGCTTTAATCTTGACGAG	829
Db	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260
QY	830	GGCCATGCATCAGAAATCCCATCCGGTTTCATTTCTACATCTTGAACCCCATGACAAC	889
Db	261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	280
QY	890	CAGAACTCAGAGTAGCTAAATCTCCATGCCCGTTTAAACACACCCGCGCAGTTTGGAGT	949
Db	281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	300
QY	950	TTCTTCCCGCGAGACGCGAGACCAATCATCTTCTGAGGCGCTTCAGCAGGAATACG	1009
Db	301	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	320
QY	1010	TTGAGCGCCCTTCATTCGGGAATTCATGAGATACGGAGGCTGCTGTTAGAGAGAAT	1069
Db	321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn	340

QY 1070 GCAGAGGTGAGCAAGAGGAGAGAGCGGAGCGGTAGTACTCGAGTGTAGAGAAC 1129
Db |||||
341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAGAGGTAGTACTCAAGTGTCAAGGACACGTTCAAGAACTTACTAAGCAGCT 1189
Db |||||
361 AsnGluGlyValIleValIleValIleValIleValIleValIleValIleValIle 380
QY 1190 AATCCGTCTTCAAGAAAGGCTCCGAGAGAGGAGAGATATACCAACCACTCAACTTG 1249
Db |||||
381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGAGGAGCGCCGATCTTCTTCAACAATTGGGAAGTTATTGAGTGAAGCCAGAC 1309
Db |||||
401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValIleValIle 420
QY 1310 AGAAGAACCCCGAGCTCAGGACCTGGACATGATCTCACCTGTGTAGAGTCAAGAA 1369
Db |||||
421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGGTTATCTCGTCTCAACAAA 1429
Db |||||
441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValIleValIleVal 460
QY 1430 GGAACCTGGAACCTTGAACTCTGCTGTGAAGAAAGAGCAACACAGAGGGAGCGG 1489
Db |||||
461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAAGAGAGGAGCAACAGCAGAGAGAGAGGAGGAGGAGTAACAGAGAGGTGCGTAGGTAC 1549
Db |||||
481 GluGluGluGluAspGluAspGluGluGluGluGluGluGluGluGluGluGluGlu 500
QY 1550 ACAGCGAGGTGAAGGAAGCGCATGTGTTTCATCATGCCAGCAGCTCATCAGTAGCCATC 1609
Db |||||
501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AAGCTTCTCCGAACTCCATCTGCTTGGTTCGCTGATCAACGCTGAAACCAACACAGA 1669
Db |||||
521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTCTCTGAGGTGATAGGACATGTATAGACCATAGACAGAGCAAGCGAGGAT 1729
Db |||||
541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGlnLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGTGCGGTGAACAAAGTTGAGAAGCTCATCAAAAACAGAGGAATCT 1789
Db |||||
561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCGTCTCGTCTCGTCTGAGAAAGAG 1849
Db |||||
581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTGAGAAAGAGATCAAGAGGAGGAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
Db |||||
601 SerProGluLysGluAspGlnGluGluGluGluGluGluGluGluGluGluGluGlu 620
QY 1910 ATTTTGAAGCTTTTAAAC 1927
Db |||||
621 IleLeuLysAlaPheAsn 626

RESULT 8
US-10-899-551-53
; Sequence 53, Application US/108999551
; Publication No. US20050063994A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Burks, A. Wesley
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Howard, Sosin B.
; APPLICANT: Bottomly, Kim H.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy
; FILE REFERENCE: 2002834-0233
; CURRENT APPLICATION NUMBER: US/10/899,551

; CURRENT FILING DATE: 2004-07-26
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 53
; LENGTH: 635
; TYPE: PRT
; ORGANISM: species Arachis hypogea
US-10-899-551-53

Alignment Scores:
Pred. No.: 1,41e-253 Length: 635
Score: 3192.00 Matches: 605
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.8% Mismatches: 0
Query Match: 89.0% Indels: 0
DB: 5 Gaps: 0

US-10-728-323-1 (1-2032) x US-10-899-551-53 (1-635)

QY 110 TCTGCAACGATGCCCAAGTCATCAGCTTACCAAGAAAGAAACAGAGAACCCCTCGCGCCAG 169
Db |||||
17 SerSerThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 36
QY 170 AGTGCTCTCAGAGTTGTCAACAGGAAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCT 229
Db |||||
37 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 56
QY 230 CCGTCACCAACGCTCAGATGATCATCTCGTCTGTCTATGATCTCGAGGACACACTGGC 289
Db |||||
57 ArgCysThrLysLeuGluLutyrAspProArgCysValTyrAspProArgGlyHisThrGly 76
QY 290 ACCACCAACCAACGTTCCCTCCAGGGAGCGGACACGCTGGCGCCCAACCCCGAGACTAC 349
Db |||||
77 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 96
QY 350 GATGATAGCCGCTCAACCCCGAAGAGAGAGGAGGAGCCGATGGGACCCAGCTGGACCG 409
Db |||||
97 AspAspAspArgArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGlyPro 116
QY 410 AGGAGAGCTGAAGAGAGAGACACTGGAGACAAACCAAGAGAGATTGGAGGCGACCAAGT 469
Db |||||
117 ArgGluArgGluArgGluLutyrArgGlnProArgGluAspTrpArgArgProSer 136
QY 470 CATCAGCAGCCACGGAATAAGGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
Db |||||
137 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluLutyrGlyThr 156
QY 530 CAGGTAGCCATGTGAGGAGAAACATCTCGGAAACACCCCTTCTACTTCCCGTCAAGG 589
Db |||||
157 ProGlySerHisValArgGluLutyrSerArgAsnAsnProPheTyrPheProSerArg 176
QY 590 CGGTTTAGCACCCGCTACGGGAACCAAAACGCTAGGATCCGGTCTCTCGAGAGGTTTGAC 649
Db |||||
177 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 196
QY 650 CAAAGGTCAAGGACGTTTCCAGAACTCCAGAACTCACCGTATTGTGCAGATCCAGGCGCAA 709
Db |||||
197 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 216
QY 710 CCTAACACTTGTCTTCCCAAGCACGCTGATGCTGATAACATCTCTGTTTATCCAGCAA 769
Db |||||
217 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 236
QY 770 GGGCAAGCCACCGTCACCGTAGCAATAGCAATAGCAATAGCAATAGCAATAGCAATAG 829
Db |||||
237 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 256
QY 830 GGCCATGCACCTCAGAAATCCCATCCCGTTTCATTTCTCTACATCTTGAACCGCCATGACA 889
Db |||||
257 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 276
QY 890 CAGAACCTCAGAGTAGTAAATCTCCATGCCCGTTTAAACACACCCCGCCAGTTTGAGAT 949
Db |||||

Db 277 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 296
QY 950 TTCTTCCCGGAGCCGAGACCAATCATCTTGCAGGGCTTCAGCGGAATACG 1009
Db 297 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 316
QY 1010 TTGAGGCGCGCTTCAATCGGAAATCAATGAGATACGAGGGTGTCTGTAGAGAAAT 1069
Db 317 LeuGluAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 336
QY 1070 GCAGGAGGTGAGCAAGAGAGAGAGCGGATGGAGTCTCGGAGTAGTGAGAAC 1129
Db 337 AlaGlyGlyGlnGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 356
QY 1130 AATGAAGGAGTGATAGTCAAAGTGTCAAAGGACACGTTCAAAGAACTTACTTAAGCACGT 1189
Db 357 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 376
QY 1190 AATCCGTCTCAAGAAAGGCTCGAAGAGAGGGAGATATCACCAACCCAACTTGTG 1249
Db 377 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 396
QY 1250 AGAAGAGGCGAGCCCATCTTTCAACAACTTTGGGAAGTTATTTCAGGTTGAAGCCAGAC 1309
Db 397 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 416
QY 1310 AGAAGAACCCCGCTTCAGGACCTGGACATGATCTCACCTGTGTAGAGATCAAGAA 1369
Db 417 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 436
QY 1370 GAGGCTTGTGATCTCCACACTTCAACTCAAAGCCATGGTTATCTGCTGCTGCAACAA 1429
Db 437 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 456
QY 1430 GGAACCTGGAACCTTGAACCTGCTGCTGTAAGAAAAGAGCAACAACAGAGGGGACCGCG 1489
Db 457 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 476
QY 1490 GAAGAAGAGGAGACGAAGACGAAGAGAGGGAAGTAAACAGAGAGTGCCTAGGTAC 1549
Db 477 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 496
QY 1550 ACAGCGAGGTGAGGAGCGATGTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
Db 497 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 516
QY 1610 AACGCTTCTCCCAACTTCCATCTGCTTGGCTTCGGTATCAACGCTGAAAACACCAACAG 1669
Db 517 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 536
QY 1670 ATCTTCTTCAGTGATAGGCAATGTGATAGACCAAGATAGAGAGCAAGCGAAGGAT 1729
Db 537 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 556
QY 1730 TTAGCATTCCTCGGTCCGTTGAACAGTTGAGAGCTCATCAAAACCAAGAGGATCT 1789
Db 557 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 576
QY 1790 CACTTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCCGCTCGTCTCGTAGAAGAG 1849
Db 577 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerProGluLysGlu 596
QY 1850 TCTCTGAGAAAGAGGATCAAGAGGAGGAGAAACCAAGAGGAGGAGGTCCACTCTTTTCA 1909
Db 597 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyLysGlyProLeuLeuSer 616
QY 1910 ATTTTGAAGCTTTTAACT 1927
Db 617 IleLeuLysAlaPheAsn 622

RESULT 9

US-09-731-221-78

; Sequence 78, Application US/09731221

; Patent No. US20020018778A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael
; TITLE OF INVENTION: Passive Desensitization
; FILE REFERENCE: 2002834-0103
; CURRENT APPLICATION NUMBER: US/09/731,221
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 78
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Arachis
; OTHER INFORMATION: Hypogaea
US-09-731-221-78

Alignment Scores:
Pred. No.: 4,76e-242 Length: 634
Score: 3052.00 Matches: 584
Percent Similarity: 96.9% Conservative: 3
Best Local Similarity: 96.4% Mismatches: 19
Query Match: 85.1% Indels: 0
DB: 3 Gaps: 0

US-10-728-323-1 (1-2032) x US-09-731-221-78 (1-634)

QY 110 TCTGCAACGCATCCCAAGTCATCACCTTACCAGAGAAACAGAGAACCCCTGCGCCAG 169
Db 16 SerSerThrHisAlaLysSerSerProTyrGlnAlaLysThrGluAsnProCysAlaGln 35
QY 170 AGTGCCTCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGATCGGAGTCT 229
Db 36 ArgCysLeuGlnSerCysGlnGlnGluProAspAlaLeuLysGlnLysAlaCysGluSer 55
QY 230 CGCTGCACCAAGCTCGAGTATGATCCTCGTGTGTCTATGATCCTCGAGACACACTGGC 289
Db 56 ArgCysThrLysLeuGluTyrAspProArgCysAlaTyrAspProArgGlyHisThrGly 75
QY 290 ACCACCAACCAAGTTTCCCTCTCAGGGAGCGGACACACTGGCGGCCAACCCGAGACTAC 349
Db 76 ThrThrAsnGlnArgSerProGlyGluAlaThrArgGlyArgGlnProGlyAspTyr 95
QY 350 GATGATGACCGCGTCAACCCCGAGAGAGAGGAGCGGATGGGAGCAGCTGACCG 409
Db 96 AspAspAlaArgArgGlnProArgAlaGluGluGlyGlyArgTrpGlyProAlaGlyPro 115
QY 410 AGGGAGCGTGAAGAGAAAGAGACTCGAGACACCAAGAGAGAGATTGGAGGCGACCAAGT 469
Db 116 ArgGluArgGluArgGluGluAspAlaArgGlnProArgGluAspTrpAlaArgProSer 135
QY 470 CATCAGCAGCCCGGAAAAATAAGGCCCGAAGAGAGAGAGAGAAACAAGAGTGGGAAACA 529
Db 136 HisGlnGlnProArgLysAlaArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 155
QY 530 CAGGTAGCCATGTGAGGAAGAAACATCTCGGAACAACCTTTCTACTTCCGTCAGG 589
Db 156 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 175
QY 590 CGGTTTTAGCACCGCTACCGGAAACCAAAACGGTAGGATCCGGGTCTCGAGAGGTTTGAC 649
Db 176 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 195
QY 650 CAAAGGTCAAGGAGTTTCAGAAATCTCCAGAAATCACCGTATTTGTGAGATCGAGGCGAAA 709
Db 196 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 215
QY 710 CCTAACACTTCTTCTTCTCCCAAGCAGCTGATGCTGATAACATCTTGTATTCCAGCAA 769
Db 216 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 235
QY 770 GGGCAAGCCACCGTAGCCGTAGCAATAGCAATAACAGAAAGAGCTTTAATCTTTGACGAG 829

Db 236 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 255
QY 830 GCCCATGCACATCCATCCGCTTCATTTCTACATCTTGAACCGCCATGACAC 889
Db 256 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 275
QY 890 CAGAACTCAGAGTACGTAATAATCTCCATGCCCGTTAAACACACACCCGCCAGTTGAGGAT 949
Db 276 GlnAsnLeuArgValAlaIleSerMetProValAsnThrProGlyGlnMetGluAsp 295
QY 950 TTCTTCCCGGAGAGCGAGACCAATCATCTTCTGACGGCTTCAGGGATACG 1009
Db 296 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheAlaArgAsnThr 315
QY 1010 TTGGAGCGCCCTTCAATCGGAATTCATAGATACGAGGGTGTCTTAGAAGAGAT 1069
Db 316 LeuGluAlaAlaPheAsnAlaGluAlaAsnGluIleArgValLeuLeuGluAsn 335
QY 1070 GCAGGAGGTGACGACAGAGAGAGCGCAGAGCGATGAGTACTCGAGTAGTGAGAAC 1129
Db 336 AlaGlyGlyGluGlnGluAlaArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn 355
QY 1130 AATGAAGAGTATAGTCAAAAGTGTCAAAGGAGCACGTTGAAGAACTTACTAAGCACGCT 1189
Db 356 AsnGluGlyValIleValIleValIleValSerLysGluHisValGluLeuThrLysHisAla 375
QY 1190 AATCCGTCTCAAGAAAGCGCTCCGAGAGAGGGAGATATCACCAACCACTCACTTG 1249
Db 376 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProAlaAsnLeu 395
QY 1250 AGAAGAGCGGAGCCCGATCTTTTAAACAACTTTGGGAAGTTATTTGAGGTGAAGCCAGAC 1309
Db 396 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuAlaGluValLysProAsp 415
QY 1310 AAGAAGAAACCCAGCTTCAGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369
Db 416 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 435
QY 1370 GGAGCTTTCATGCTCCACACTCAACTCAAGGCCATGTTATCGTCTGCTCAACAA 1429
Db 436 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 455
QY 1430 GGAACCTGGAACCTTCAACTCTGGCTGTAAAGAAAGAGACAAACAGAGGGGAGCGCG 1489
Db 456 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 475
QY 1490 GAAGAAAGAGGAGCAGAACCGAAGAGAGGAGGAAAGTAAACAGAGAGTGCCTAGGTAC 1549
Db 476 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgAlaTyr 495
QY 1550 ACAGCGAGTTGAGGAAGCGATGTGTTTCATCTCCAGCAGCTCATCCAGTACCATC 1609
Db 496 ThrAlaArgLeuLysGlyAspValPheIleMetProAlaAlaHisProValAlaIle 515
QY 1610 AACGCTTCTCCGAACCTCCATCTGCTTGGCTTCGGTATCAACGCTCAAAACAAACACAGA 1669
Db 516 AsnAlaSerSerGluLeuAlaLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg 535
QY 1670 ATCTCTCTTCGAGGTGATAGGCAATGTGTATAGACCATGATAGAGAAGCAAGGAT 1729
Db 536 IlePheLeuAlaGlyAspAlaAspAsnValIleAspGlnIleGluLysGlnAlaLysAla 555
QY 1730 TTAGCATTCCTGGGTCCGGTGNACCAAGTTGAGAACTCATCAAAACCAAGAGGATCT 1789
Db 556 LeuAlaAlaProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 575
QY 1790 CACTTGTGAGTCTCGTCTCTCAATCTCAATCTCAATCTCCGTCGCTCTCGAGAAGAG 1849
Db 576 HisPheValAlaAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 595
QY 1850 TCTCTCGAGAAAGAGATCAAGAGGAGGAAACCAAGGAGGAGGTCCTCCTCTTCA 1909

Db 596 SerProGluLysGluAspGlnGluAsnGlnGlyGlyLysGlyProLeuLeuSer 615
QY 1910 ATTTTGAAGGCTTTTAAAC 1927
Db 616 IleLeuLysAlaPheAsn 621
RESULT 10
US-09-331-631A-21
; Sequence 21, Application US/09331631A
; Patent No. US20020168392A1
; GENERAL INFORMATION:
; APPLICANT: Manners, John M.
; APPLICANT: Marcus, John Paul
; APPLICANT: Goulter, Kenneth C.
; APPLICANT: Green, Jodie L.
; TITLE OF INVENTION: ANTIMICROBIAL PROTEINS
; FILE REFERENCE: CULAN23 001APC
; CURRENT APPLICATION NUMBER: US/09/331,631A
; PRIOR FILING DATE: 1999-06-21
; PRIOR APPLICATION NUMBER: PCT/AU97/00874
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: AU PO 4275
; PRIOR FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Peanut
; US-09-331-631A-21
Alignment Scores:
Pred. No.: 3,78e-241 Length: 614
Score: 3041.00 Matches: 595
Percent Similarity: 96.0% Conservative: 9
Best Local Similarity: 94.6% Mismatches: 7
Query Match: 84.8% Indels: 18
DB: 3 Gaps: 7
US-10-728-323-1 (1-2032) x US-09-331-631A-21 (1-614)

QY 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTGCTAGGATCCTTGTCTGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGCATCCCAAGTCATCCTTACCAGAGAGAAACACAGAACCCCTCGGCCAG 169
Db 21 SerAlaThrGlnAlaLys---SerProTyr---ArgLysThrGluAsnProCysAlaGln 38
QY 170 AGGTCCCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAACAAAGGCATGCGAGTCT 229
Db 39 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 58
QY 230 CGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCTTCGAGGACACACTGC 289
Db 59 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAsp-----ThrGly 74
QY 290 ACCACCAACCAACGTTCCCTCCAGGGAGCGGACACGTCGGCCGCCAACCCCGAGACTAC 349
Db 75 AlaThrAsnGlnArgHisProProGlyGluArgThrArgGlnProGlyAspTyr 94
QY 350 GATGATGACCGCGCTCAACCCCGAAGAGAGAGGAGGCGCGATGGGGACAGCTGGACCG 409
Db 95 AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGluPro 114
QY 410 AGGGAGCGTGAAGAGAGAACAGACTGGAGACAAACCAAGAGAGATTGGAGGGACCAAGT 469
Db 115 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 134
QY 470 CATCAGCAGCCACCGAATAATAGGCCCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
Db 135 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 154

```
QY 530 CCAGGTAGCATGTGAGGGAAGAAACATCTCGAACAACCCCTTCTACTCCGTCGAG 589
Db 155 ProGlySerGluValArgGluGluThrSerArgAsnAsnProPheThrPheProSerArg 174
QY 590 CGGTTTACGACCCCGCTACCGGAACCAAAACGGTAGGATCCGGTCCCTCGACAGGTTTGAC 649
Db 175 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 194
QY 650 CAAAGTCAAGGCAGTTTCAGAAATCTCCAGATCTCCAGATTCACCGTATTGTGAGATCGAGGCCAA 709
Db 195 GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg 214
QY 710 CCTAACACATCTGTTCTTCCCAAGCACGCTGATGCTGATAACATCTTGTATCCAGCAA 769
Db 215 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 234
QY 770 GGGCAAGCCACCGTAGCCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTTCACGAG 829
Db 235 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 254
QY 830 GGCATGCACTCAGAAATCCCATCCCGTTTCATTCCTACATCTTGAAACCGCCATGACAAC 889
Db 255 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 274
QY 890 CAGAACTCAGAGTAGCTAAATCTCCATGCCCGTTAAACACACCCCGCCAGTTTGAGGAT 949
Db 275 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 294
QY 950 TTCTTCCCGCGAGCAGCCAGACCAATCATCTACTTGCAGGCTTCAGCAGGAATACG 1009
Db 295 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 314
QY 1010 TTGAGCGCCGCTTCAATGCGGAATTCATAGATACGAGGGTGTGTTAGAGAAGAT 1069
Db 315 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 334
QY 1070 GCAGGAGCTCAGCAAGGAGAGCGGCAGCGGATGGAGTACTCGGAGTAGTGAGAAC 1129
Db 335 AlaGlyGlyGluGlnGluArgGlyGlnArgArgSerThrArgSerSerAsp--- 353
QY 1130 AATGAAGGAGTGATAGTCAAAGTGTCAAAGCAGCTTGAAGAACTTACTAAGCACGCT 1189
Db 354 AsnGluGlyValIleValLysValSerLysGluHisValGlnGluLeuThrLysHisAla 373
QY 1190 AATCCGTTCTCAAAGAAAGCTCCGAAGAGAGGGAGATATCAACCAACCAATCAACTTG 1249
Db 374 LysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu 392
QY 1250 AGAAGAGCGGAGCCCGATCTTTCTAACAACCTTTGGGAAGTTATTTCAGGTGAAGCCAGAC 1309
Db 393 ArgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValLysProAsp 412
QY 1310 AAGAAGAACCCCGAGCTTCAGGACCTCGACATGCTCACCTGTGTAGAGATCAAGAAGAA 1369
Db 413 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 432
QY 1370 GAGAGCTTGATGTCCTCCACATCTCAACTCAAGGCCATGTTATCGTGTGCTCAACAAA 1429
Db 433 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 452
QY 1430 GGAACCTGGAACCTTGAACCTCGTGGCTGTAAGAAAGAGCAACAACAGAGGGGACCGCG 1489
Db 453 GlyThrGlyAsnLeuGluLeuAlaValArgLysGluGlnGlnArgGlyArgGly 472
QY 1490 GAA-----GAAGAGGAGGACGAAGACGAAGAAGAGGAGGGAAGTAAACAGAGGGTG 1540
Db 473 GluGlnGlnLrpGluGluGluGluAspGluGluGluGlySerAsnArgGluVal 492
QY 1541 CGTAGGTACACAGCGAGGTTGAAGGAAGGCGATGTGTTTCATCATGCCAGCAGCTCATCCA 1600
Db 493 ArgArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisPro 512
QY 1601 GTAGCCATCAACGCTTCCCTCCGAACTCCATCTCTGCTGGCTTCGGTATCAACGCTGAAAC 1660
```

```
Db 513 ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn 532
QY 1661 AACCAACAGAAATCTCTTCGACAGTATAAGGCAATGTGATAGACACAGATAGAGAAGCAA 1720
Db 533 AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln 552
QY 1721 GCGAAGGATTTAGCATTCCTCGGTGGGTGAACAAGTTGAGAAGCTCATCAAAACCCAG 1780
Db 553 AlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGln 572
QY 1781 AAGGAATCTCACTTTGTGAGTGCTCGTCTCTCAATCTCAATCTCAATCTCCGTCGCTCCT 1840
Db 573 ArgGluSerHisPheValSerAlaArgProGlnSerGlnSerProSer----- 588
QY 1841 GAGAAAGAGTCTCTCGAGAAAGAGATCAAGAGGAGGAGAAACCAAGAGGAGGAGGTCCA 1900
Db 589 -----SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyGlyPro 605
QY 1901 CTCCTTTCAAATTTTGAAGGCTTTTAAC 1927
Db 606 LeuLeuSerIleLeuLysAlaPheAsn 614

RESULT 11
US-09-847-208-27
; Sequence 27, Application US/09847208
; Publication No. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; TITLE OF INVENTION: IGE-MEDIATED ALLERGIC DISEASES
; FILE REFERENCE: UC67.002A
; CURRENT APPLICATION NUMBER: US/09/847,208
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Arachis hypogaea (Peanut)
US-09-847-208-27

Alignment Scores:
Pred. No.: 3,78e-241 Length: 614
Score: 3041.00 Matches: 595
Percent Similarity: 96.0% Conservative: 9
Best Local Similarity: 94.6% Mismatches: 7
Query Match: 84.8% Indels: 18
DB: 3 Gaps: 7

US-10-728-323-1 (1-2032) x US-09-847-208-27 (1-614)
QY 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTGCTAGGATCCTTGTCTCGTGGTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGCATGCCCAAGTCATCACCTTACCAGAGAAACACAGAGAACCCCTCGCCCCAG 169
Db 21 SerAlaThrGlnAlaLys--SerProTyr---ArgLysThrGluAsnProCysAlaGln 38
QY 170 AGGTGCTTCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAAGGCATCGAGTCT 229
Db 39 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 58
QY 230 CGCTGCACCAAGCTCCAGTATGATCCTCGTTGTGTCTATCATCTCCAGGACACACTGGC 289
Db 59 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAsp-----ThrGly 74
QY 290 ACCACCAACCAAGTTCCCTCCAGGGAGCGGACACGTGGCCGCCCAACCCGAGACTAC 349
Db 75 AlaThrAsnGlnArgHisProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 94
```

```
QY 350 GATGATGACCGCGTCAACCCGAGAGAGAGAGGCGCGATGGGACACGAGTGGACCG 409
Db |||
95 AspAspAspArgArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGluPro 114
QY 410 AGGAGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 469
Db |||
115 ArgGluArgGluArgGluArgGluArgGluArgGluArgGluArgGluArgGluArgGlu 134
QY 470 CATCAGCAGCCGCGGAAATAGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
Db |||
135 HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGluGluGluGluGluGlu 154
QY 530 CCAGGTAGTCATGTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 589
Db |||
155 ProGlySerGluValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 174
QY 590 CGGTTTAGCACCGCTACCGGAAACCAAAACGGTAGGATCCGGGTCTCGAGAGGTTTGAC 649
Db |||
175 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 194
QY 650 CAAAGCTCAAGGCAGTTTCAGAAATCTCCAGAAATCCAGATCCAGATTCGAGATCCGAGGCCARA 709
Db |||
195 GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg 214
QY 710 CCTAACACACTCTGTTCTCCCAAGCACGCTGTATGATAACATCTCTGTTATCCAGCAA 769
Db |||
215 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 234
QY 770 GGGCAAGCACCGTGACCGTAGCAATGCAATACAGAAAGAGCTTTAATCTTGACGAG 829
Db |||
235 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 254
QY 830 GCGCATGCACTCAGATCCATCCGCTTCATCTCTACATCTTGACCGCCATGACACAC 889
Db |||
255 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 274
QY 890 CAGAACCTCAGAGTACGTAATCTCCATGCCCGTTAAACACACCCCGCCAGTTTGAAGAT 949
Db |||
275 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 294
QY 950 TTCTTCCCGGAGCAGCAGCAGCAATCATCTCTGAGCGGCTTCAGCGAGGATACG 1009
Db |||
295 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 314
QY 1010 TTGGAGCGCGCTTCAATCGGAATTCATGATACGAGCGGCTGTAGAGAGAT 1069
Db |||
315 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 334
QY 1070 GCAGGAGGTGAGCAAGAGAGAGAGCGGATGGAGTACTCGGAGTAGTGAGAAC 1129
Db |||
335 AlaGlyGlyGluGlnGluArgGlyGlnArgArgSerThrArgSerSerAsp--- 353
QY 1130 AATGAAGAGTATGATCTAAAGTGTCAAGGAGCAGCTTGAAGAACTTCTAAGCAGCT 1189
Db |||
354 AsnGluGlyValIleValIleValIleValSerLysGluHisValGlnGluLeuThrLysHisAla 373
QY 1190 AATCCGTCTCAAGAAAGCTCCGAGAGAGGAGATATCACCAACCAATCACTG 1249
Db |||
374 LysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu 392
QY 1250 AGAAGAGGCGAGCGCCGATCTTTCTAAACACTTTGGGAAGTTATTTGAGGTGAAGCCAGAC 1309
Db |||
393 ArgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValLysProAsp 412
QY 1310 AAGAAGAACCCCGCTTCAGGACCTGGACATGATCTCACTGTGTAGAGATCAAGAA 1369
Db |||
413 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 432
QY 1370 GAGGCTTTGATGCTCCACACTTCACTCAAGGCCATGGTTATCGTCTGCTCAACAA 1429
Db |||
433 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 452
```

```
QY 1430 GGAACCTGAAACCTTTGAACTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGACGGCGG 1489
Db |||
453 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 472
QY 1490 GNA-----GAAGAGGAGACCAAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1540
Db |||
473 GluGlnGluTrpGluGluGluAspGluGluGluGluGluGluGluGluGluGluVal 492
QY 1541 CQTAGGTACACAGCCAGGTGGAAGGAGCGATGTGTTTCATCGCCAGCAGCTCATCCA 1600
Db |||
493 ArgArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisPro 512
QY 1601 STAGCCATCAACGCTTCTCCGAACCTCATCTGCTTGGCTTCGGTATCAACGCTGAAAC 1660
Db |||
513 ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn 532
QY 1661 AACCAACAGAACTTCTTCAGGTCATTAAGCAATGTGATAGACCAAGATAGAGAACAA 1720
Db |||
533 AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln 552
QY 1721 GCGAAGGATTTAGCATTCCTCGGTGGGTGAACAAAGTTGAGAAGCTCATCAAAACCCAG 1780
Db |||
553 AlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuLysAsnGln 572
QY 1781 AAGGAATCTCACTTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCAATCTCTCTCT 1840
Db |||
573 ArgGluSerHisPheValSerAlaArgProGlnSerGlnSerProSer----- 588
QY 1841 GAGAAGAGTCTCTCGAAGAGAGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1900
Db |||
589 -----SerProGluLysGluAspGlnGluGluGluGluGluGluGluGluGlyLysGlyPro 605
QY 1901 CTCCTTTCAATTTTGAAGGCTTTTAAAC 1927
Db |||
606 LeuLeuSerIleLeuLysAlaPheAsn 614

RESULT 12
US-10-147-095-21
; Sequence 21, Application US/10147095
; Publication No. US20030171274A1
; GENERAL INFORMATION:
; APPLICANT: Manners, John M.
; APPLICANT: Marcus, John Paul
; APPLICANT: Goulter, Kenneth C.
; APPLICANT: Green, Jodie L.
; TITLE OF INVENTION: ANTIMICROBIAL PROTEINS
; FILE REFERENCE: CULLN23.001APC
; CURRENT APPLICATION NUMBER: US/10/147,095
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US/09/331.631A
; PRIOR FILING DATE: 1999-06-21
; PRIOR APPLICATION NUMBER: PCT/AU97/00874
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: AU PO 4275
; PRIOR FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Peanut
US-10-147-095-21

Alignment Scores:
Pred. No.: 3,78e-241 Length: 614
Score: 3041.00 Matches: 595
Percent Similarity: 96.0% Conservative: 9
Best Local Similarity: 94.6% Mismatches: 7
Query Match: 84.8% Indels: 18
DB: Gaps: 7

US-10-728-323-1 (1-2032) x US-10-147-095-21 (1-614)
```


QY 50 ATGAGAGGAGGCTTTCTCCACTGATGCTGTTGCTTAGGATCCTTGTCTGCTTCCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGCATGCCAAGTCATCACCTTACCAAGAAAGAAACAGAGAACCCCTGCGCCACG 169
Db 21 SerAlaThrGlnAlaLys---SerProTyr--ArgLysThrGluAsnProCysAlaGln 38
QY 170 AGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCGATGCGAGTCT 229
Db 39 ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 58
QY 230 CGTGTACCAACAGTCCAGATGATGATCCTGTTGTGTTCTATGATCCTCGAGACACACTGGC 289
Db 59 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAsp-----ThrGly 74
QY 290 ACCACCAACCAACGTTCCCTCCAGGGGAGCGGACACAGTGGCGGCCCAACCCGAGACTAC 349
Db 75 AlaThrAsnGlnArgHisProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 94
QY 350 GATGATGACCGCGTCAACCCCAAGAGAGGAAGGAGCGCGATGGGACCAAGTGGAGCGACCACT 469
Db 95 AspAspArgArgGlnProArgGluGlyArgGlyArgGlyArgGlyProAlaGluPro 114
QY 410 AGGAGCGTCAAGAGAGAGAGACTGGAGACCAACCAAGAGAGACTGGAGCGACCAACT 469
Db 115 ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgProSer 134
QY 470 CATCAGCAGCCAGCAAAATAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
Db 135 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 154
QY 530 CCAGGTAGCCATGTGAGGGAAGAAACAATCTCGGAACAACCCCTTTACTTCCCGTCAAGG 589
Db 155 ProGlySerGluValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 174
QY 590 CGGTTTAGCACCGCTACGGGACCAACCAACGGTAGGATCCGGTCTCGCAGAGGTTTGAC 649
Db 175 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 194
QY 650 CAAAGTCAAGGAGGATTCAGAAATCTCCAGAAATCCAGTATTGTGCAGATCGAGGCGCAA 709
Db 195 GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg 214
QY 710 CCTAACACTTGTGTTCTCCAGCAGCAGCTGATGATAACATCTGTTGTTATCCAGCAA 769
Db 215 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 234
QY 770 GGGCAACCCACCGTAGCCGTAGCAATGGCAATAACAGAAAGAGCTTAACTCTTGACGAG 829
Db 235 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 254
QY 830 GGCATCACTCAGAAATCCCATCCGTTTCATTTCTTACATCTTGAACCCGCCATGACAAC 889
Db 255 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 274
QY 890 CAGAACTCAGTAGTAAATCTCCATGCCCGTTAAACACACCCGCGCCAGTTTGAGGAT 949
Db 275 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 294
QY 950 TTCTTCCCGGAGCAGCCGAGCAATCATCTTACTTGCAGGCTTCAGCAGGAATACG 1009
Db 295 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 314
QY 1010 TTGAGGCGCGCTTCAATGGGGAATTCATCAGATACGAGGGTGTGTTTAGAAGAGAAT 1069
Db 315 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 334
QY 1070 GCAGGAGGTAGCAAGAGAGAGAGGCGCAGGCGATGGAGTACTCGGAGTAGTAGAAC 1129
Db 335 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgSerThrArgSerSerAsp--- 353
QY 1130 AATGAAGGAGTGATAGTCAAAAGTGTCAAAGGAGCAGCCTTGAAGAACTTACTAAGCAGCT 1189

Db 354 AsnGluGlyValIleValLysValSerLysGluHisValGlnGlnLeuThrLysHisAla 373
QY 1190 AAATCCGCTTCAAAGAAAGGCTCCGAAAGAGAGGAGATATCACCAACCCAAATCAACTTG 1249
Db 374 LysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu 392
QY 1250 ACAGAGAGGCGAGCCGATCTTCTAACACTTTGGGAAGTATTATTTGAGGTGAGGAGCAGAC 1309
Db 393 ArgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValLysProAsp 412
QY 1310 AAGAAGAAACCCCGAGCTTCAGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAGAA 1369
Db 413 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 432
QY 1370 GGAGCTTGTATGCTCCACACTTCAACTCAAAAGGCGATGTTATTCGTCTCGTCAACAAA 1429
Db 433 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 452
QY 1430 GGAACCTGNAACCTTGAACCTCGTCTGTAAAGAAAGAGCAACACAGAGGGGAGCGCG 1489
Db 453 GlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnArgGlyArgArg 472
QY 1490 GAA-----GAAGAGGAGCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGGTTG 1540
Db 473 GluGlnGluTrpGluGluGluGluAspGluGluGluGlySerAsnArgGluVal 492
QY 1541 CTAAGTACACAGCAGGTTGAAGAGGCGATGTGTTTCATCATGCCAGCAGCTCATCCA 1600
Db 493 ArgArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaHisPro 512
QY 1601 GTAGCCATCAACGCTTCTCCGAACTCCATCTGCTTGGCTTCGGTATCAACCTGAAAAAC 1660
Db 513 ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn 532
QY 1661 AACCACAGAAATCTTCTTCAGAGTATAGAGCAATGTGATAGACAGATAGAGAGCAA 1720
Db 533 AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln 552
QY 1721 GCGAAGGATTAGATTCCTCCGCGGTGAGCAAGTTGAGAGAGCTCATCAAAACCCAG 1780
Db 553 AlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGln 572
QY 1781 AAGGAATCTCACTTGTGAGTCTGCTCTCAATCTCAATCTCAATCTCAATCTCCGTCCTCT 1840
Db 573 ArgGluSerHisPheValSerAlaArgProGlnSerGlnSerProSer----- 588
QY 1841 GAGAAAGAGTCTCTCAGAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1900
Db 589 -----SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyPro 605
QY 1901 CTCCTTCAATTTGAAGCTTTTAAAC 1927
Db 606 LeuLeuSerIleLeuLysAlaPheAsn 614
RESULT 13
US-10-100-303A-8
; Sequence 8, Application US/10100303A
; Publication No. US20030202980A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; FILE REFERENCE: 2002834-0166
; CURRENT APPLICATION NUMBER: US/10/100,303A
; CURRENT FILING DATE: 2002-03-18
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 614
; TYPE: PRF
; ORGANISM: Arachis hypogaea, Prot/(Nucleo-Ara h 1
US-10-100-303A-8

Alignment Scores: 3.78e-241 Length: 614
Pred. No.: 3041.00 Matches: 595
Score: 96.0% Conservatives: 9
Percent Similarity: 94.6% Mismatches: 18
Best Local Similarity: 84.8% Indels: 7
Query Match: 4 Gaps: 7
DB: 1

US-10-728-323-1 (1-2032) x US-10-100-303A-8 (1-614)

QY 50 ATGAGAGGAGGTTTCTCCACTGATGCTGTGGTATAGGATCCTTGTCTGCTTCAGTT 109
|||||
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACCGCATGCCAAGTCATCCTTACCAGAGAAACAGAGAACCCCTGCCGCCAG 169
|||||
Db 21 SerAlaThrGlnAlaLys---SerProTyr---ArgLysThrGlnAsnProCysAlaGln 38
QY 170 AGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCT 229
|||||
Db 39 ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 58
QY 230 CGTGCACCAAGCTCAGTATGATCTCTGTTGTCTATGATCTTCGAGGACACACTGGC 289
|||||
Db 59 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAsp-----ThrGly 74
QY 290 ACCACCAACCAAGTTCCTCCAGGGAGCGGACGTCGGCGCCCAACCCGAGACTAC 349
|||||
Db 75 AlaThrAsnGlnArgHisProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 94
QY 350 GATGATGACCGCGCTCAACCCCGAAGAGAGGAGGCGCGATGGGACACAGCTGGACCG 409
|||||
Db 95 AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGluPro 114
QY 410 AGGGAGCGTGAAGAGAGAAGACTGGAGACAAACCAAGAGAAGATTGGAGGCGACCAAGT 469
|||||
Db 115 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 134
QY 470 CATCAGCAGCCAGGAATAGGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
|||||
Db 135 HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGluGluGluGluGluGlu 154
QY 530 CCAGGTAGCATGTGAGGGAAGAAACATCTCGAAACAAACCTTTCTACTTCCCGTCAAGG 589
|||||
Db 155 ProGlySerGluValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 174
QY 590 CGGTTTAGCACCGCTACGGGAACCAAAACGGTAGATCCGGGTCTTCAGAGGTTTGAC 649
|||||
Db 175 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 194
QY 650 CAAAGGTCAAGGCAGTTTCAGAAATCTCCAGAAATCCCGTATGTCAGATCCGAGGCCAAA 709
|||||
Db 195 GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg 214
QY 710 CTTAACACTCTGTGTTCTTCCCAAGCAGCTGATGCTGTATACATCTTGTATCCAGCAA 769
|||||
Db 215 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 234
QY 770 GGGCAAGCCACCGTAGCCGTAGCAATATGCAATATACAGAAAGAGCTTTAATCTTCAGCAG 829
|||||
Db 235 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 254
QY 830 GGCCATGCACTCAGAAATCCATCCGTTTCATTTCTCATCATCTTGAACCGCCATCACAAC 889
|||||
Db 255 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 274
QY 890 CAGAACTCAGAGTAGTAAAATCTCCATGCCCGTTAACACACCCGGCCAGTTTGGAGAT 949
|||||
Db 275 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 294
QY 950 TTCTTCCCGGAGGAGCGGAGCAATCATCTTCTGACGGCTTCAGAGGATACG 1009
|||||

RESULT 14

US-10-100-303A-55

; Sequence 55, Application US/10100303A

; Publication No. US20030202980A1

; GENERAL INFORMATION:

; APPLICANT: Caplan, et al.

; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction

Db 295 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerSerArgAsnThr 314
QY 1010 TTGGAGGCGCGCTTCAATCGCAATTCAGATACGAGGCTCTCTTTAGAACAGAAAT 1069
|||||
Db 315 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn 334
|||||
QY 1070 GCAGAGGTGAGCAAGAGGAGGAGGCGATGGATGATCTCGGAGTAGTGAGAAC 1129
|||||
Db 335 AlaGlyGlyGluGlnGluArgGlyGlnArgArgSerThrArgSerSerAsp--- 353
QY 1130 AATGAGGAGTGTAGTCAAAAGTGTCAAGGAGCAGCTTGAAGACTTACTTAAGCACCT 1189
|||||
Db 354 AsnGluGlyValIleValLysValSerLysGluHisValGlnGluLeuThrLysHisAla 373
|||||
QY 1190 AAATCCGCTCTCAAGAAAGGCTCCGAAAGAGGAGATATCACCAACCCCAATCAACTTG 1249
|||||
Db 374 LysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu 392
|||||
QY 1250 AGAGAAGCGAGCGCGATCTTTTCTTAACAACCTTTGGGAAGTTATTGTAGGTGAAGCCAGC 1309
|||||
Db 393 ArgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValLysProAsp 412
|||||
QY 1310 AAGAAGAACCCCGAGCTTCAGGACCTGCACATGATCTCACCTGTGTAGAGATCAAGAA 1369
|||||
Db 413 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 432
|||||
QY 1370 GGAGCTTTGATGTCTCCACACTTCAACTCAAGGCGCATGTTATCGTCTCGTCAACAAA 1429
|||||
Db 433 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 452
QY 1430 GGAATCGAAACCTTGAACTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGACGGCG 1489
|||||
Db 453 GlyThrGlyAsnLeuLeuValAlaValArgLysGluGlnGlnArgGlyArg 472
QY 1490 GAA-----GAAGAGGAGACGAACGAGAGAGGAGGAGGAGTAACAGAGAGGTG 1540
|||||
Db 473 GluGlnGluTrpGluGluGluAspGluGluGluGluGluGlySerAsnArgGluVal 492
QY 1541 CGTAGGTACACAGCGAGGTTGAAGGAGCGCATGTGTTTCATCATCCAGCAGCTCATCCA 1600
|||||
Db 493 ArgArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisPro 512
QY 1601 GTAGCCATCAACGCTTCTCCGAACTCATCTGTTGGCTTCGGTATCAACGCTGAAAAAC 1660
|||||
Db 513 ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn 532
QY 1661 AACCAACAGAACTTCTTCAGAGTGTAAAGACAAATGTGTAGACCAATAGAGAGCA 1720
|||||
Db 533 AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln 552
QY 1721 GCGAAGGATTTAGCATTCCTCGGTGGGTGAAACAAAGTTGAGAACTCATCAAAAAACCAG 1780
|||||
Db 553 AlalysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGln 572
QY 1781 AAGGAATCTACTTTGTGAGTGTCTCTCAATCTCAATCTCAATCTCCGCTGCTCTCT 1840
|||||
Db 573 ArgGluSerHisPheValSerAlaArgProGlnSerGlnSerProSer----- 588
QY 1841 GAGAAAGAGTCTCTGAGAAAGAGGATCAAGAGGAGGAAACCAAGAGGAGGAGGAGGTC 1900
|||||
Db 589 -----SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyPro 605
QY 1901 CTCCTTTTCAATTTTGAAGGCTTTTAAAC 1927
|||||
Db 606 LeuLeuSerIleLeuLysAlaPheAsn 614

```

; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0166
; CURRENT APPLICATION NUMBER: US/10/100,303A
; CURRENT FILING DATE: 2002-03-18
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Arachis hypogaea, Prot/Nucleo-Ara h 1
US-10-100-303A-55

Alignment Scores:
Pred. No.: 111e-105 Length: 268
Score: 1394.00 Matches: 268
Percent Similarity: 100.0% Conservativeness: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 38.9% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-1 (1-2032) x US-10-100-303A-55 (1-268)

QY 485 AAAATAAGGCCCAAGGAAGAGAGAAACAAGAGTGGGGAACACACAGGTAGCCATGTG 544
Db 1 LysileargProGluGlyArgGluGlyGluGlnGluTrpGlyThrProGlySerHisVal 20

QY 545 AGGGAAGAAACATCTCGGAACAACCCCTTCTACTTCCCGTCAAGGCGGTTTACGACCCCGC 604
Db 21 ArgGluGluThrSerArgAsnAsnProPheTy:PheProSerArgPheSerThrArg 40

QY 605 TAGGGNACCAAAACGGTAGGATCCGGTCTCTCAGAGGTTTCAACAAAGGTCAAGGCAG 664
Db 41 TyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAspGlnArgSerArgGln 60

QY 665 TTTCAGAATCTTCAGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTAACACTCTTGT 724
Db 61 PheGlnAsnLeuGlnAsnHisA:G:IleValGlnIleGluAlaLysProAsnThrLeuVal 80

QY 725 CTTCCCAAGCACGCTGATGCTGATAACATCCTTGTATTATCCAGAAAGGCCAACCCGCTG 784
Db 81 LeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGlyGlnAlaThrVal 100

QY 785 ACCGTAGCAATCGCAATACAGAAAGAGCTTTAACTTCACGAGGCCCATCGACATCAGA 844
Db 101 ThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGluGlyHisAlaLeuArg 120

QY 845 ATCCCATCCCGTTTCATTTCCTACATCTTGAAACCGCCATCACCAACCAACCTCAGAGTA 904
Db 121 IleProSerGlyPheIleSerTyIleLeuAsnArgHisAspAsnGlnAsnLeuArgVal 140

QY 905 GCTAAATCTCCATGCCCCGTATACACACCCCGCCAGTTTGAAGATTTCCTCCCGCGAGC 964
Db 141 AlaLysIleSerMetProValAsnThrProGlyGlnPheGluAspPhePheProAlaSer 160

QY 965 AGCCGAGACAATCATCTACTTCAGGGCTTCAGCAGGAATACGTTGGAGGCGCGCTTC 1024
Db 161 SerArgAspGlnSerSerTyLeuGlnGlyPheSerArgAsnThrLeuGluAlaAlaPhe 180

QY 1025 AATCCGGAATTCATAGATACGAGGGTGCTGTGTAGAAGAGATGCGAGAGGTGAGCAA 1084
Db 181 AsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsnAlaGlyGlyGluGln 200

QY 1085 GAGAGAGAGGGCAGGCGGATGGAGTACTCGGAGTAGTGAGAACATGAAGAGGTGATA 1144
Db 201 GluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsnAsnGluGlyValIle 220

QY 1145 GTCAAACTGTCAAGGAGCAGGTGAGAGACTTACTAAGCACCGCTAAATCCGCTCAAAG 1204
Db 221 ValIysValSerIysGluHisValGluGluLeuThrLysHisAlaLysSerValSerLys 240

QY 1205 AAAGGCTCCGAAGAGGGAGATATCACCAACCCAACTCACTTGAGAGAGCGAGCC 1264
Db 241 LysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeuArgGluGlyGluPro 260

QY 1265 GATCTTTCTAACAACTTTGGGAAG 1288
Db 261 AspLeuSerAsnAsnPheGlyLys 268

RESULT 15
US-10-424-599-153195
; Sequence 153195, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 153195
; LENGTH: 605
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_10935C.1.pap
US-10-424-599-153195

Alignment Scores:
Pred. No.: 2.59e-96 Length: 605
Score: 1282.00 Matches: 285
Percent Similarity: 59.4% Conservativeness: 108
Best Local Similarity: 43.1% Mismatches: 173
Query Match: 35.8% Indels: 96
DB: 4 Gaps: 18

US-10-728-323-1 (1-2032) x US-10-424-599-153195 (1-605)

QY 50 ATGAGAGGAGGGTTTCTCCACTGATGTGTTGCTTAGGATCTTCTCGGCTTCAGTT 109
Db 2 MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyLeuValPheLeuAlaSerVal 19

QY 110 TCTGCCAACGCATCCCAAGTCATCACCTTACCAGAGAAACAGAACCCCTCGGCCAG 169
Db 20 SerValSerPheGlyIleAla-----TyrTrpGluLys---GluAsnProLysHisAsn 36

QY 170 AGGTGCTCTCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATCGAGTCT 229
Db 37 LysCysLeuGlnSerCysAsnSerGluArgAspSerTy:ArgAsnGlnAlaCysHisAla 56

QY 230 CGCTGCACCAAGTCTGAGTATGAT-----CCTCGTTGT 262
Db 57 ArgCysAsnLeuLeuLysValGluLysGluGluCysGluGluGlyGluIleProArg-- 75

QY 263 GTCTATGATCCTCGAGGACACACTGGCCACCACCAACGTTCCCTCCAGGGGAGCGG 322
Db 76 -----ProArgProArg-----ProGlnHisProGlu 84

QY 323 ACACGTGGCGCCCAACCCGGAGACTACGATGATGAC----- 358
Db 85 ArgGluProGlnGlnProGlyGluLysGluGluAspGluAspGluGlnProArgProIle 104

QY 359 -----CGCGTCAACCCCGAAGAGAG-----GAGAGAGGCCGA 391
Db 105 ProPheProArgProGlnProArgGlnGluGluHisGluGlnArgGluGluGlnGlu 124

QY 392 TGG-----GGACCAGCTGGAGCCGAGGAGCGTGAAGAGAGAA 430
Db 125 TrpProArgLysGluGluLysArgGlyGluLysGlySerGluGluGluAspGluAspGlu 144

QY 431 GACTGGAGACAACCAAGAAAGAT-----TGGAGCGCAACCAAGTCAATCAGCAGCCCG 484
Db 145 AspGluGluGlnAspGluArgGlnPheProPheProArgProHisGlnLysGluGlu 164
```


THIS PAGE LEFT BLANK

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 21:23:02 ; Search time 60.8696 Seconds
(without alignments)
4400.310 Million cell updates/sec

Title: US-10-728-323-1
Perfect score: 3586
Sequence: 1 aataatcatatattcatc.....cgttgtgctgtttcttccc 2032

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-O=/abs/ABSWEB spool/US10728323/runat_15052006_172130_22366/app_query.fasta.1
-DB=A_Geneseq -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p
-USER=US10728323 @CGN 1 1 605 @runat_15052006_172130_22366 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq 21:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*
9: Geneseq2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3286	91.6	626	2 AAY15244	Aay15244 Peanut al
2	3286	91.6	626	2 AAY25657	Aay25657 Peanut al
3	3286	91.6	626	7 ABUS2412	Abus2412 Peanut Ar
4	3286	91.6	626	7 ADC34906	Adc34906 Peanut al
5	3286	91.6	626	9 ADV97601	Adv97601 Peanut Ar
6	3282	91.5	626	4 AAU04706	Aau04706 Anaphylac
7	3282	91.5	626	6 ABUS2570	Abus2570 Peanut Ar
8	3280	91.5	626	6 ABUS2568	Abus2568 Peanut Ar
9	3280	91.5	626	6 ABUS2571	Abus2571 Peanut Ar

10	3280	91.5	626	6 ABUS2574	Abus2574 Peanut Ar
11	3280	91.5	626	6 ABUS2573	Abus2573 Peanut Ar
12	3279	91.4	626	8 ADM12053	Adm12053 Arachis h
13	3278	91.4	626	6 ABUS2569	Abus2569 Peanut Ar
14	3278	91.4	626	6 ABUS2572	Abus2572 Peanut Ar
15	3274	91.3	626	2 AAW22150	Aaw22150 Peanut al
16	3269.5	91.2	625	7 ADG27464	Adg27464 Peanut Ar
17	3251	90.7	626	8 ADO38271	Ado38271 Peanut al
18	3052	85.1	634	3 AAB33599	Aab33599 Modified
19	3052	85.1	634	4 AAU04709	Aau04709 Modified
20	3052	85.1	634	4 AAU05034	Aau05034 Modified
21	3041	84.8	614	2 AAW22149	Aaw22149 Peanut al
22	3041	84.8	614	2 AAW62834	Aaw62834 Arachis h
23	3041	84.8	614	6 ABUS2413	Abus2413 Peanut Ar
24	3041	84.8	614	7 ADG27465	Adg27465 Peanut Ar
25	3034	84.6	614	9 ADV97606	Adv97606 Peanut Ar
26	2130	59.4	415	2 AAY40913	Aay40913 Ara h 1 a
27	1700	47.4	335	9 ADV97604	Adv97604 Peanut Ar
28	1281	35.7	605	2 AAW62838	Aaw62838 Glycine m
29	1281	35.7	605	7 ADG27564	Adg27564 Soybean B
30	1276	35.6	605	2 AAY40999	Aay40999 Soybean b
31	1274.5	35.5	623	8 ADT58763	Adt58763 Plant pol
32	1190	33.2	559	5 ABG71270	Abg71270 Glycine m
33	1180.5	32.9	543	5 ABG71271	Abg71271 Glycine m
34	1165	32.5	417	5 ABG71269	Abg71269 Glycine m
35	1157	32.3	417	6 ABP97239	Abp97239 Mature be
36	1157	32.3	425	6 ABP97238	Abp97238 FLAG-tag
37	1156	32.2	439	6 ABP97235	Abp97235 Glycine m
38	1117.5	31.2	390	6 ABP97236	Abp97236 Glycine m
39	874	24.4	324	8 ADX75042	Adx75042 Plant ful
40	867	24.2	344	8 ADX89039	Adx89039 Plant ful
41	865.5	24.1	666	2 AAW62829	Aaw62829 Macadamia
42	863.5	24.1	625	2 AAW62830	Aaw62830 Macadamia
43	849.5	23.7	666	2 AAW62828	Aaw62828 Macadamia
44	828	23.1	316	8 ADX79896	Adx79896 Plant ful
45	828	23.1	316	8 ADX74539	Adx74539 Plant ful

ALIGNMENTS

RESULT 1	
AAY15244	AAY15244 standard; protein; 626 AA.
XX	XX
XX	AC AAY15244;
XX	AC
DT	17-OCT-2003 (revised)
DT	09-NOV-1999 (first entry)
XX	XX
DE	Peanut allergen, Ara h 1, amino acid sequence.
XX	XX
KW	allergy; immune response; transgenic; allergen; epitope;
KW	immunoglobulin E; Ig E; binding site; peanut.
XX	XX
OS	Arachis hypogaea.
XX	XX
PN	WO9938978-A1.
PD	05-AUG-1999.
XX	XX
PF	29-JAN-1999; 99WO-US002031.
XX	XX
PR	31-JAN-1998; 98US-0073283P.
PR	13-FEB-1998; 98US-0074590P.
PR	13-FEB-1998; 98US-0074624P.
PR	13-FEB-1998; 98US-0074633P.
PR	27-AUG-1998; 98US-00141220.
XX	XX
PA	(UYAR-) UNIV ARKANSAS.
PA	(UYNY) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.
XX	(SOSI/) SOSIN H.
XX	Sosin H, Bannon GA, Burks AW, Sampson HA;
PI	PI

XX WPI; 1999-479189/40.
 DR N-PSDB; AAZ06382.
 XX Modified allergen with reduced IgE binding, useful for treating e.g. allergies.
 PT
 PT
 XX
 PS Disclosure; Page 35-37; 46pp; English.

XX This is the amino acid sequence of the Ara h 1 protein from *Arachis hypogaea*. The Ara h 1 protein has 23 IgE (immunoglobulin E) binding epitopes, four of which are immunodominant (AAV15247, AAV15249, AAV15250 and AAV15263). By modifying the IgE binding sites the ability of the allergen to provoke an immune response is downregulated. The epitopes of the IgE binding sites can therefore be modified in genetically engineered plants and animals to elicit less of an allergic response. (Updated on 17 -OCT-2003 to standardise OS field)

XX Sequence 626 AA;

Alignment Scores:
 Pred. No.: 1,599-304 Length: 626
 Score: 3286.00 Matches: 626
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 91.6% Indels: 0
 DB: 2 Gaps: 0

US-10-728-323-1 (1-2032) x AAV15244 (1-626)

QY 50 ATGAGAGGAGGTTTCTCCACTGATGCTGTTCTAGGATCCTGTTCTCGGCTTCAGTT 109
 Db 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
 QY 110 TCTGCAACGATGCCAAGTCATCACCTTACAGAGAAACACAGAACCCCTCGCCCGAG 169
 Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
 QY 170 AGGTGCTCCAGAGTTGTTCACAGGAACCGGATGACTTGAAGCAAAAGGCATCGAGTCT 229
 Db 41 ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60
 QY 230 CGGTGACCAAGCTCGAGTATGATCCTCGTTGTCTATGATCCTCGAGACACACTGGC 289
 Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
 QY 290 ACCACCAACCAACGTTCCCTCCAGGGGAGGACAGCTGGCCGCCCAACCCGGAGACTAC 349
 Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
 QY 350 GATGATGACCCCGTCAACCCGAGAGAGAGAGGAGCCGATGGGACCCAGCTGGACCG 409
 Db 101 AspAspAspArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGlyPro 120
 QY 410 AGGAGCGTCAAGAGAGAGAGCTGGAGACACCAAGAGAGATTGGAGCGCACCAAGT 469
 Db 121 ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
 QY 470 CATCAGCAGCCACGAAATAAGCCCGAAGAGAGAGAGAGAGAAAGAGTGGGAAACA 529
 Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTyrGlyThr 160
 QY 530 CCAGGTAGCCATGTGAGGGAGAAACATCTCGGAACCAACCTTTCTACTTCCCGTCAAGG 589
 Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
 QY 590 CGGTTTAGCACCCCTACGGGAACCAAAACGGTAGGATCCGGTCTCGCAGAGGTTTGAC 649
 Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
 QY 650 CAAAGGTCAAGCGAGTTTCAGAAATCTCCAGAATCACCGTATTTGTGAGATCGAGGCCAAA 709
 Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220

QY 710 CCTAACACTCTTGTTCCTCCCAAGCACGCTGATGATAACATCCTTGTATCCAGCAA 769
 Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
 QY 770 GGGCAAGCCACCGTACCGTAGCAATGGCAATACAGAAAGAGCTTTAATCTTGACGAG 829
 Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
 QY 830 GGCCATGCATCAGAAATCCCATCCGGTTTCATTTCTACATCTTGAACCCCATGACAAC 889
 Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
 QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCCCGTTAAACACACCCGCCAGTTGGAGT 949
 Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
 QY 950 TTCTCCCGCGAGCAGCAGACCAATCATCTACTTCAGCGGCTTCAGCAGGAATACG 1009
 Db 301 PhePheProAlaSerSerArgaspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
 QY 1010 TTGGAGCCCGCTTCAATCGGAAATTCATGAGATACGAGGGTGTCTTTAGAAGAAT 1069
 Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 340
 QY 1070 GCAGAGGTGACCAAGAGAGAGAGCGGATGAGTACTCGGAGTAGTGAGAAC 1129
 Db 341 AlaGlyGlyGlnGlnGluArgGlyGlnArgTyrSerThrArgSerSerGluAsn 360
 QY 1130 AATGAAGAGTAGTAGTCAAGTGTCAAAGAGACAGTTGAAGAACTTACTAAGCACGCT 1189
 Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
 QY 1190 AATCCGCTCAAGAAAGGCTCCGAAGAGAGGGAGATATCACCAACCAATCAACTTG 1249
 Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
 QY 1250 AGAAGAGGAGCCCGATCTTCTTCAACTTTGGAGACTTATTTGAGTGGAGCCAGAC 1309
 Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
 QY 1310 AGAAGAAACCCCGAGCTTCAGGACCTTGACATGATCCTCACCTGTGTAGAGATCAAGAA 1369
 Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu 440
 QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAAGGCCATGGTTATCTGCTGTCACAAA 1429
 Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
 QY 1430 GGAACCTGGAACCTTCAACTCGTGGCTGAAGAAAAGACCAACACAGAGGGGACCGCG 1489
 Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgTyr 480
 QY 1490 GAAGAGAGAGGAGCAGAACGAAAGAGAGAGGAGGAAAGTAACAGAGAGGTGCTAGGTAC 1549
 Db 481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgTyr 500
 QY 1550 ACAGCAGGTTGAAGAGGCGATGTTCATCATCTCCAGCAGCTCATCAGTAGCCATC 1609
 Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
 QY 1610 AACGCTTCTCCGAACCTCCATCTGCTTGGCTTCGGTATCAACGCTCAAAACCAACACAGA 1669
 Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg 540
 QY 1670 ATCTTCTTCGAGGTGATAGGACAAATGTGATAGACCATAGAGAGCAAGCAAGGAT 1729
 Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
 QY 1730 TTAGCATTCCTCGGTCGGGTGAACAAAGTTGAGAACTCATCAAAACCAAGAGGATCT 1789
 Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580

Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleAArgValLeuLeuGluAsn 340
QY 1070 GCAGGAGGTGAGCAAGAGAGAGAGCGGATGAGTACTCGAGTAGTGAGAAC 1129
Db 341 AlaGlyGlyGluGlnGluAArgGlyGlnAArgTrpSerThrArgSerGluAsn 360
QY 1130 AATGAAGGAGTGTAGTCAAAAGTGTCAAAGGACAGCTTGAAGAACTTACTAAGCAGCT 1189
Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluLeuThrLysHisAla 380
QY 1190 AATCCGTCTCAAGAAGCGCTCCGAGAGAGGAGATATCACCAACCCACTCACTTG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProfileAsnLeu 400
QY 1250 AGAAGAAGCGGAGCCGATCTTTCAACACTTTGGGAAGTTATTTCAGGTGAAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAACCCCGAGCTTCAGGACCTGGACATGCTCACCTGTGTAGAGATCAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 CGAGCTTGTGCTCCACACTTCAACTCAAAGCCATGGTTATCGTCTCGTCAACAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAACCTGGAACCTTGAACCTCGTGGCTGTAAAGAAAAGAGCAACAGAGGGGACGGCG 1489
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyAArg 480
QY 1490 GAAGAAGAGGAGCAGAAAGCAGAAAGAGAGGAGGAGTAAACAGAGAGTGCCTAGGTAC 1549
Db 481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValAArgArgTyr 500
QY 1550 ACAGCGAGGTGAAGAGCGGATGTCTTCATCATCCAGCAGCTCATCCAGTAGCCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AAGCTTCTCCGAACCTTCAACTCTGCTTGGCTTCAGCTCAACGCTGAAACACACACAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCCTCCAGGTGATAGGCAATGTGATAGACCATAGACATAGAGAAGCGAAGGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGTCGGGTGAACAGTTGAGAGCTCATCAAAACACAGAGGAATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCCGTCGTCTCCCTGAGAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerProGluLysGlu 600
QY 1850 TCTCTGAGAAAGAGGATCAAGAGGAGGAAACCAAGAGGAGGAGGTCACCTCTTTCA 1909
Db 601 SerProGluLysGluAspGlnGluGluGluGluGluGluGluGluGluGluGluGlu 620
QY 1910 ATTTTGAAGGCTTTTAAAC 1927
Db 621 IleLeuLysAlaPheAsn 626

RESULT 3
ID ABU52412
XX ABU52412 standard; protein; 626 AA.
AC ABU52412;
XX
XX 10-MAR-2003 (first entry)
XX Peanut Ara h1 protein from cDNA clone P41b.
XX

KW Peanut; allergy; Ara h1; Ara h2; Ara h3; IgE binding site;
KW anaphylactic food allergen; anti-allergenic; vaccine; wound healing.
XX
XX Arachis hypogaea.
XX WO200274250-A2.
XX
XX 26-SEP-2002.
XX
XX 18-MAR-2002; 2002WO-US009108.
XX
XX 16-MAR-2001; 2001US-0276822P.
XX 18-MAR-2002; 2002US-00276822.
XX (PANA-) PANACEA PHARM.
XX Caplan M, Sosin H, Sampson H, Bannan GA, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabjohn PA, Shin DS, Stanley JS;
XX WPI; 2003-018765/01.
XX N-PSDB; ABX70603.
XX
XX New modified anaphylactic food allergen, useful for preventing or
XX treating allergic reactions associated with e.g. anaphylactic allergens.
XX
XX Claim 27; Fig 13; 300pp; English.
XX
XX The invention relates to a modified anaphylactic food allergen has an
XX amino acid sequence that is substantially identical to that of natural
XX anaphylactic food allergen, except for a cysteine residue that has been
XX modified so that it cannot participate in the disulphide bond. The
XX modification may also comprise mutation of the IgE binding sites to
XX reduce allergenicity. Also included are: (1) a method of making a
XX modified anaphylactic food allergen; (2) a nucleotide molecule encoding
XX or for causing a site specific mutation in the modified anaphylactic food
XX allergen; (3) a transgenic plant or animal expressing the modified
XX anaphylactic food allergen; (4) a method of treating an individual by
XX reducing the clinical response to a natural anaphylactic food allergen;
XX and an isolated fragment of peanut allergen Ara h 1. The modified
XX anaphylactic food allergen is useful for preventing or treating allergic
XX reactions associated with any natural allergen such as food, insect,
XX rubber or preferably anaphylactic allergens. It is also useful for
XX treating wounds in mammals such as bovine, canine, feline, caprine,
XX ovine, porcine, murine or equine species. The present sequence is a
XX peanut allergen (e.g. Ara h1, h2 or h3)
XX
XX Sequence 626 AA;

Alignment Scores:
Pred. No.: 1.59e-304 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.6% Indels: 0
DB: 6 Gaps: 0
US-10-728-323-1 (1-2032) x ABU52412 (1-626)

QY 50 ATGAGAGGGAGGGTTCCTCCACTGATGCTGTGTAGGATCCTGTCTGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGCATGCCAAGTCATCACCTTACCAGAGAAACAGAGAACCCCTCGCCCGAG 169
Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGGTGCTCCAGAGTTGTCAACAGGACCGGATGACTTGAAGCAAAAGGCATCGAGTCT 229
Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGGC 289

Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACAAAGCTTCCCTCCAGGGGAGCGACACGTCGGCCGCAACCCGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCGTCAACCCGGAAGAGAGGAGCGCGATGGGACCCAGCTGGACCG 409
Db 101 AspAspArgArgGlnProArgGluGluGlyArgTrpGlyProAlaGlyPro 120
QY 410 AGGAGCGTCAAGAGAGAGACTGGAGACACCAAGAGAAATTTGGAGCGCACCAAGT 469
Db 121 ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgProSer 140
QY 470 CATCAGCAGCCACGGAATAAGCCCGAAGGAGAGAGAGAAACAAGAGTGGGAAACA 529
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGlnGluTrpGlyThr 160
QY 530 CCAGGTAGCCATGTCAGGGAAGAAACATCTCGGAACAACCCCTTTCTACTTCCCGTCAAG 589
Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCGCTACGGGAACCAAAACGGTAGGATCCGGTCTCTGCAGAGGTTTGAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGGTCAGGCGAGTTTCAGAATCTCCAGAATCCAGCAATCTGTGCAGATCCGAGGCCAAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTTCTTCCCAAGCACGCTGATGCTGATAACATCCTGTGTATCCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGCAGACCACTGACCGTAGCAATGGCAATPAAACAGAAAGCTTTAATCTTGACGAG 829
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCATGCACTCAGATCCCATCCGCTTTCATTTCTTCTACATCTTGACCGCCATGACAAC 889
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACTCAGAGTAGTAAATCTCATGCCCGTTTAAACACACCCGCCAGTTTGGAGGAT 949
Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCAGCAGCCGAGACCAATCATCTACTTCAGGGCTTCAGCAGGAATACG 1009
Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGGCGCCTTCAATCGGAATTCATAGATACGAGGGTCTGTTAGAAGAGAAAT 1069
Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGGTGAGCAAGAGGAGAGCGGAGCGGATGGAGTACTCGGAGTAGTGAGAAC 1129
Db 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGGAGTATAGTCAAAGTGTCAAAGGAGCAGCTTGAAGAACTTACTAAGCAGCT 1189
Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCGCTCTCAAGAAAGCTCCGAGAGAGGGAGATATCACCAACCAATCACTTG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGAGCGGAGCCCATCTTTCTAAACACTTTGGGAAGTTATTGAGGTGAAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGACCCCGCTTACGACCTGGACATGATCTCACCTGTGTAGAGATCAAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440

QY 1370 GGAGCTTTTGATGCTCCACACTTCAACTCAAAAGGCCATGGTTATCGTCTCGTCAACAAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
QY 1430 GGAACCTGGAACCTTGAACTCGTGTGTAAGAAAAGACAAACAAGAGGGACCGCGG 1489
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAAGAGAGGAGGAGCAAGACCAAGAGAGGAGGAGTAACAGAGAGGTGCGTAGGTAC 1549
Db 481 GluGluGluGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTTCAGGAAGGCCATGTGTTTCATCTATGCCAGCAGCTCATCCAGTAGCCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGTTCTCTCCGAATCTCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACACAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg 540
QY 1670 ATCTTCTTGCAGGTGATNAAGGACAAATCTGTATAGACCATAGAGAACCAAGCGAGGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGTTCGGGTCAACAAGTTGAGAAAGCTCATCAAAAACCAAGGAATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTCTCGTCTCTCAATCTCAATCTCAATCTCCTGCTCTCCTGAGAAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCCTGAGAAAGAGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
Db 601 SerProGluLysGluAspGlnGluGluGluGlnGluGluGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTTAAAC 1927
Db 621 IleLeuLysAlaPheAsn 626
RESULT 4
ADC34906
ID ADC34906 standard; protein; 626 AA.
XX ADC34906;
AC ADC34906;
XX 18-DEC-2003 (first entry)
XX Peanut allergen Ara h 1.
XX allergen; antigen; hyporesponsive; desensitisation; immunomodulator;
KW gene therapy; peanut.
XX Arachis hypogaea.
XX WO2003047618-A2.
XX 12-JUN-2003.
XX 05-DEC-2002; 2002WO-GB005548.
XX 05-DEC-2001; 2001US-0338385P.
XX (CIRC-) CIRCASSIA LTD.
XX Larche M, Ledger PW;
XX WPI; 2003-523267/49.
Desensitizing an individual to a selected polypeptide antigen comprises
PT administering a composition containing polypeptide antigens in an amount
PT that generates a state of hyporesponsiveness to the antigen to allow

desensitization.

Disclosure; Page 40; 57pp; English.

The invention relates to a novel method for desensitising an individual to a selected polypeptide antigen. The method comprises administering a composition that contains polypeptide antigens in an amount that generates in the individual a state of hyporesponsiveness to the antigen to allow desensitisation to one or more polypeptide antigens. The method of the invention has immunomodulator activity, and may have a use in gene therapy. The composition and method are useful in manufacturing a medicament for desensitising an individual to a selected polypeptide antigen or for generating in the individual a state of hyporesponsiveness to the antigen to allow desensitisation to one or more polypeptide antigens. The present sequence is used in the exemplification of the invention.

Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCCTGACAAGAGCATCAAGAGGAGGAACCAAGAGGAGGAGGTCCTCTTCA 1909
Db 601 SerProGluLysGluAspGlnGluGluGluAaGlnGlyGlyGlySerProLeuLeuSer 620
QY 1910 ATTTGAAGCTTTTAAC 1927
Db 621 IleLeuLysAlaPheAsn 626

RESULT 5
ADV97601
ID ADV97601 standard; protein; 626 AA.
AC ADV97601;
XX DT 10-MAR-2005 (first entry)
XX Peanut Ara h I allergen alpha P17 seed storage protein.
XX allergen; atopic dermatitis; diagnostic; immunotherapy;
KW immediate type hypersensitivity; Ara h I; dermatological; antiallergic;
KW immunosuppressive; antibody engineering; seed storage protein.
XX Arachis hypogaea.
XX US6835824-B1.
XX PD 28-DEC-2004.
XX PF 13-NOV-1998; 98US-00191593.
XX PR 29-DEC-1995; 95US-0009455P.
XX PR 23-SEP-1996; 96US-00717933.
XX PA (UYAR-) UNIV ARKANSAS.
XX PI Burks AW, Stanley JS, Bannon GA, Cockrell G, Helm RM;
XX WPI: 2005-045982/05.
XX DR N-PSDB; ADV97537, ADV97600.

PT New isolated nucleic acid encoding peanut allergen Ara h 1, useful for
PT obtaining diagnostics with optimal concentrations of each allergen, for
PT developing panels of mixtures of recombinant allergens, and in
PT immunotherapy.

XX Disclosure; Fig 16; 14lpp; English.

XX This invention relates to a novel isolated nucleic acid molecule encoding
CC the peanut allergen Ara h I that has been identified in patients with
CC atopic dermatitis. Specifically, it refers to a monoclonal antibody
CC enzyme-linked immunosorbent assay (ELISA) used to detect and quantify the
CC peanut allergens, where the monoclonal antibodies have specificity for a
CC selected peanut allergen antigen. The present invention also describes
CC hybridomas that produce two-site monoclonal antibodies specific for a
CC peanut allergen that can be used in an ELISA to detect and determine the
CC concentration of a specific peanut allergen in a food product or in food
CC processing. Furthermore, it provides an expression vector and the
CC isolated nucleic acid or its fraction can include a diagnostic label.
CC Accordingly, these nucleic acids are useful for obtaining diagnostics
CC with optimal concentrations of each allergen or for developing panels of
CC mixtures of large numbers of recombinant allergens such that it can be
CC used in immunotherapy for the treatment of food hypersensitivity
CC reactions. This polypeptide sequence is the peanut Ara h I allergen
CC protein (P17 clone) of the invention.

XX Sequence 626 AA;

Alignment Scores: 1.59e-304 Length: 626
Pred. No.: 3286.00 Matches: 626
Score:

Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.6% Indels: 0
DB: 9 Gaps: 0
US-10-728-323-1 (1-2032) x ADV97601 (1-626)
QY 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTGTGCTAGGATCCTTGTCTCGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGCATGCCAAGTCATCCTTACAGAAAGAAAACAGAGAACCCCTGCGCCAG 169
Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGGTGCTCCAGAGTTGTCAACAGAACCGGATGACTTGAAGCAAAAGCATGCCAGTCT 229
Db 41 ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGC 289
Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAACGTTCCCTCCAGGGAGCGGACACGTGGCCGCCAACCCCGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCCGCTCAACCCCGAGAGAGAGAGAGAGCCGATGGGACCACTGGACCG 409
Db 101 AspAspAspArgArgGlnProArgGluGluGlyArgGlyProAlaGlyPro 120
QY 410 AGGAGCGTGAAAGAGAAAGACTGGAGACAAACCAAGAGAGAGATTGGAGCGCACCACT 469
Db 121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
QY 470 CATCAGCAGCACGGAAAAATAAGCCCGAAGAGAGAGAGAGAGAGAGAGTGGGGAACA 529
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
QY 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCCCTTCTACTTCCCGTCAAGG 589
Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCCGCTAGCGGAACCAAAACCGTAGGATCCGGGTCTCTCGAGAGTTTGAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGTCAAGCGAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTGTTTCTTCCCAAGCACGCTGATGCTGATAACATCCTTGTATTCCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTGACCGTAGCAATAGCAATAGCAATAGCAATAGCAATAGCAATAG 829
Db 241 GlyGlnAlaThrValThrValAlaAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCCATGACTCAGAAATCCCATCCCGTTTCATTTCTCATCTTGAACCGCCATGACAAC 889
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGATAGTCTTAAATCTCCATGCCCGTTAAACACACCCCGCCAGTTTGGAGAT 949
Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGGCGAGCGAGCGAGACCAATCATCTACTTTCAGGGGCTTCAGCAGGAATACG 1009
Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGAGGCGCGCTTCAATCGGAATTCAATGAGATACGAGGGTCTCTGTAGAGAGAT 1069

Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleAArgArgValLeuLeuGluAsn 340
QY 1070 GCAGGAGGTGAGCAAGAGGAGAGCGGATGGAGTACTCGGAGTAGTGAGAAC 1129
Db 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTTPSerThrArgSerGluAsn 360
QY 1130 AATGAAGAGTAGTAGTCAAAAGTGTCAAGGAGCAGCTTGAAGAACTTACTAAGCAGCT 1189
Db 361 AsnGluGlyValIleValIysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AAATCCGTCTCAAGAAAGCTCCGAGAGAGGAGAGATATACCAACCACTCAACTTG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAGAAAGGCGAGCCGATCTTTTAAACAATTTGGGAAGTATTTCAGGTGAAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAACCCAGCTTCAGGACCTGGACATGCTACCTGTGTAGAGATCAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCTCGTCGTCACAAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAACCTGGAAACCTTGAACCTCGTGTGAAGAAAGAGCAACAAGAGGGGACGGCG 1489
Db 461 GlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnArgGlyArgG 480
QY 1490 GAAGAGAGGAGACGAGAGCAAGAGAGAGGAGGAGTAACAGAGGTGCTAGGTAC 1549
Db 481 GluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgGlyr 500
QY 1550 ACAGCGAGTTGAAGGAGCGATGTCTTCATGCAGCAGCTATCCAGTAGGCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleWetProAlaHisProValAlaIle 520
QY 1610 AAGCTTTCCTCCGAACCTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACACACAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCCTCCAGTGATAGGACAATGCTGATAGACCAGATAGAGAGCGAAGGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTGGCTCGGTGAACAGTTTGAGAGGCTCATCAAAACACAGAGGAGTCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTCTGAGTGTCTGCTCTCAATCTCAATCTCAATCTCGTCTCTGAGAAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerProGluLysGlu 600
QY 1850 TCTCCTCAGAAAGAGGATCAGAGGAGGAGAAACCAAGGAGGAGGCTCCACTCTTCA 1909
Db 601 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTTAAAC 1927
Db 621 IleLeuLysAlaPheAsn 626

RESULT 6

AAU04706
ID AAU04706 standard; protein; 626 AA.
XX
AC AAU04706;
XX
DT 23-OCT-2001 (first entry)
XX
DE Anaphylactic antigen Ara h 1.
XX
KW Ara h 1; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;
Alignment Scores: 3.84e-304 Length: 626
Pred. No.: 3282.00 Matches: 625
Score: 99.8% Conservative: 0
Percent Similarity: 99.8% Mismatches: 1
Best Local Similarity: 91.5% Indels: 0
Query Match: 4 Gaps: 0
DB: 0
US-10-728-323-1 (1-2032) x AAU04706 (1-626)
QY 50 ATCAGAGGGAGGGTTCCTCCACTGATGCTGTTCTAGGATCCTTGTCTGGTTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGATGCCAAGTCATCCCTTACCAGAAACACAGAACCCCTCGGCCAG 169
Db 21 SerAlaThrHisAlaLysSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40

KW allergy; mast cell; basophil; mouse.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT Misc-difference 285
FT /note= "Encoded by gta"
XX
PN WO200140264-A2.
XX
PD 07-JUN-2001.
XX
PF 06-DEC-2000; 2000WO-US033124.
XX
PR 06-DEC-1999; 99US-00455294.
PR 23-JUN-2000; 2000US-0213765P.
PR 27-SEP-2000; 2000US-0235797P.
XX
PA (PANA-) PANACEA PHARM LLC.
PA (UYAR-) UNIV ARKANSAS.
PA (MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.
XX
PI Bannan GA, Burks WA, Caplan MJ, Sampson H, Sosin H;
XX
DR WPI; 2001-381378/40.
DR N-PSDB; AAS08537.

Antigenic fragments useful for reducing anaphylactic risk and reducing the severity and/or number of allergic symptoms in individuals sensitive to antigens, have reduced ability to bind Immunoglobulin E.

Claim 7; Fig 9; 100pp; English.

The sequence represents the amino acid sequence of anaphylactic antigen Ara h 1. Ara h 1 is an anaphylactic antigen (A), which was used to design antigenic peptides having a reduced ability to bind immunoglobulin E (IgE) as compared with the intact (A), or having a sequence substantially identical to a portion of sequence of an antigen that includes at least one IgE binding site, where at least one IgE binding site of the peptide is altered. The antigenic peptides are used in a composition which is useful for reducing risk or severity of allergic reaction to an antigen. This is done by identifying an individual at risk of allergic reaction to an antigen by identifying prior display of allergic symptoms when exposed to the antigen, or a familial relationship with an individual who previously displayed allergic symptoms when exposed to the antigen. Following this an antigen-specific IgE present on one or more mast cells or basophils in the individual's serum is identified. The individual is then contacted with a peptide corresponding to a portion of the antigen, which is selected, formulated, and delivered so that binding of the peptide to antigen-specific IgE is reduced as compared with IgE binding of intact antigen. The composition is also useful for treating and preventing allergic reactions

Sequence 626 AA;

QY 170 AGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATCGAGTCT 229
Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGGTGCACCAAGCTCAGATATGATCTCGTTGTGCTATGATCTCTCAGGACACACTGGC 289
Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAAGTTCCTCCAGGGGAGCGGACACAGCTGGCGCCCAACCCGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATACCGCGCTCAACCCGAGAGAGAGGAGCGCGATGGGACACAGCTGACCG 409
Db 101 AspAspArgArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGlyPro 120
QY 410 AGGGAGCTGAAGAGAAGACTGGAGACAACCAAGAGAAGATTGGAGGGGACCAAGT 469
Db 121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
QY 470 CATCAGCAGCCAGGAAATAGGCCCGGAAGAGAGAGAGAAACAAGAGTGGGGAACA 529
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
QY 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGAACAACCCCTTCTACTTCCCGTCAAGG 589
Db 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCGCTACGGGAACCAAAACGGTAGGATCGGGTCTCGAGAGTTTGAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGTCAAGGCAGTTTCAGAAATCTCCAGATCCACCGTATTGTGCAGATCCAGGCCAAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTTGTTCTCCCAAGCACGCTGATGCTGATAACATCTCTGTTATCCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTAGCCGTAGCAATGGCAATACAGAAAGAGCTTTAATCTTGACGAG 829
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCATGCACTCAGAAATCCATCCCGGTTTCATTCTCATCTTGAACCGCCATGACAAC 889
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACTCAGAGTAGCTAAAATCTCCATGCCCGTTAAACACACCCCGCCAGTTGAGGAT 949
Db 281 GlnAsnLeuArgAlaAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGGAGCAGCGAGACCAATCATCTACTTTCAGGGCTTCAGCAGGAATACG 1009
Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGCGCGCTTCAATCGGAATTCATAGATACGAGGCTGCTGTTAGAAGAGAT 1069
Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGGTGACGACAGAGAGAGCGGAGCGGATGAGTACTCGGAGTAGTGAGAAC 1129
Db 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTipSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGGAGTAGTAGTCAAAAGTGTCAAAGGACGCTTGAAGAACTTACTAAGCAGCT 1189
Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluLeuLeuThrLysHisAla 380
QY 1190 AATCCGCTCTCAAAGAAAGCTCCGAAGAAGAGGGAGATATCACCAACCCCAATCAACTTG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGAGCGGAGCCCGATCTTCTTCAACAACTTTGGGAAGTTATTTGAGGTGAAGCCAGAC 1309

Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAGAAGACCCACCTTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluLeuLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAAGGCGCATGGTTATCGTCGTCGTCACAAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
QY 1430 GGAACCTGGAACCTTGAACCTCGTGTGTAAAGAGAAAGACAACAACAGAGGGGACGGCGG 1489
Db 461 GlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnGlnArgGlyArgArg 480
QY 1490 GAAGAAGAGGAGGACGAAGACGAAGAGAGGAGGAAGTAACAGAGAGAGGTGCGTAGGTAC 1549
Db 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTGAAGGAAGCGGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGAACCTCATCTGCTTGGCTTCGGTATCAACGCTGAAAAACAACACACA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTGCAGGTGATAAGGACAAATGTGATAGACCATAGACAGAACGCAAGGAGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTTCCCTGGTCCGGTCAACAAGTTGAGAAAGCTCATCAAAAACAGAAAGAACT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCCGCTGCTCTCTCTCAGAAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTCTGAGAAGAGGATCAAGAGGAGGAAACCAAGGAGGAGGAGGCTCCACTCTCTTCA 1909
Db 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyGlyGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTTAAAC 1927
Db 621 IleLeuLysAlaPheAsn 626
RESULT 7
ABU52570
ID ABU52570 standard; protein; 626 AA.
XX AC ABU52570;
XX XX
DT 10-MAR-2003 (first entry)
XX DE Peanut Ara h1 mutant V72A.
XX KW Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;
KW KW mutein; anaphylactic food allergen; antiallergenic; vaccine;
XX KW wound healing.
OS Homo sapiens.
XX XX
PN WO200274250-A2.
XX PD 26-SEP-2002.
XX PF 18-MAR-2002; 2002WO-US009108.
XX PR 16-MAR-2001; 2001US-0276822P.
XX PR 18-MAR-2002; 2002US-00276822.
XX XX
PA (PANA-) PANACEA PHARM.

Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
Rabjohn PA, Shin DS, Stanley JS;
WPI; 2003-018765/01.

New modified anaphylactic food allergen, useful for preventing or
treating allergic reactions associated with e.g. anaphylactic allergens.

Example 5; Page; 30pp; English.

The invention relates to a modified anaphylactic food allergen has an
amino acid sequence that is substantially identical to that of natural
anaphylactic food allergen, except for a cysteine residue that has been
modified so that it cannot participate in the disulphide bond. The
modification may also comprise mutation of the IGE binding sites to
reduce allergenicity. Also included are: (1) a method of making a
modified anaphylactic food allergen; (2) a nucleotide molecule encoding
or for causing a site specific mutation in the modified anaphylactic food
allergen; (3) a transgenic plant or animal expressing the modified
anaphylactic food allergen; (4) a method of treating an individual by
reducing the clinical response to a natural anaphylactic food allergen;
and an isolated fragment of peanut allergen Axa h 1. The modified
anaphylactic food allergen is useful for preventing or treating allergic
reactions associated with any natural allergen such as food, insect,
rubber or preferably anaphylactic allergens. It is also useful for
treating wounds in mammals such as bovine, canine, feline, caprine, and
ovine, porcine, murine or equine species. The present sequence is a food
allergen, mutated to alter its IGE binding characteristics. Note: The
present sequence is not shown in the specification but was created by the
indexer using information provided in the specification

Sequence 626 AA;

Alignment Scores:		
Pred. No.:	3,846-304	Length:
Score:	3822.00	Matches:
Percent Similarity:	99.8%	Conservative:
Best Local Similarity:	99.8%	Mismatches:
Query Match:	91.5%	Indels:
DB:	6	Gaps:
US-10-728-323-1 (1-2032)	X	ABUS2570 (1-626)

Qy	50	ATGAGGGAGGGTTCTTCCACTGATGCTGTGTCTAGGATCCTTGTCTCGGCTTCAGTT	109
Db	1	MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal	20
Qy	110	TCTGCCAACGCATGCCAAGTCATCACCTTACCAGAGAAACAGAGAACCCCTCGGCCGAG	169
Db	21	SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln	40
Qy	170	AGTGGCTCTCAGAGTTGTCAAACAGAAACCGGATGACTTGAAGCAAAAGGATCGAGTCT	229
Db	41	ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer	60
Qy	230	CGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCTCTCAGAGACACACTGGC	289
Db	61	ArgCysThrLysLeuGlnLysArgProArgCysAlaTyrAspProArgGlyHisThrGly	80
Qy	290	ACCACCAACCAAGTTCCCTCCAGGGGACGGACACGTGGCGCCCAACCCGGAGACTAC	349
Db	81	ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	100
Qy	350	GATGATGACCGCCGTCAAACCCCAAGAGAGAAAGGCGCATGGGACACCAAGTGGACCG	409
Db	101	AspAspAspArgGlnProArgGlnGluGlyArgTdpGlyProAlaGlyPro	120
Qy	410	AGGGAGCGTGAAGAAGAAGACTGGAGACAACCAAGAGAAGATTGGAGCGCACCAAGT	469
Db	121	ArgGluArgGluArgGluArgTdpArgGlnProArgGluAspTdpArgTdpProSer	140

QY 770 GGCACGCCACCGTAGCCGTAGCAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAG 829
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 GlyGlnAlaThrValThrValAlaLaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY GGCATGCACTCAGAAATCCCATCCGGTTTCATTTCTTACATCTTGAAACCGCCATGACAAC 889
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrlleLeuAsnArgHisAspAsn 280
QY CAGAACTCAGAGTAGTAAATCTCCATGCCCGTTAAACACACCGGCCCATGTTGAGGAT 949
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
281 GlnAsnLeuArgValAlaIlyIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY TTCTTCCCGCAGCAGCGAGACCAATCATCTACTTGCAGGCTTCAGCAGGAATACG 1009
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 PhePheProAlaSerArgAspGlnSerSerTyrlleuGlnGlyPheSerArgAsnThr 320
QY TTGGAGGCCCCCTTCAATGCGGAATTCATGAGATACGGAGGCTGCTGTAGAGAGAAT 1069
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 340
QY GCAGGAGGTGAGCAAGAGAGAGAGCGCAGCGATGGAGTACTGGAGTAGTGAGAAC 1129
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY AATGAAGGAGTGATAGTCAAAGTGTCAAAGGAGCAGCTTGAAGAACTTACTAAGCAGCT 1189
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
361 AsnGluGlyValIleValIlyValSerLysGlnHisValGluGlnLeuThrLysHisAla 380
QY AAATCCGCTTCAAAGAAAGGCTCCGAAGAGAGAGGAGATATCAACCAACCAATCAACTTG 1249
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY AGAAGAGCGGAGCCCGATCTTTCTAACAACCTTTGGGAAGTTATTGAGGTGAAGCAGAC 1309
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY AAGAAGAACCCCGACTTCAGGACCTCGGACATGATGCTACCTGCTGAGAGATCAAGAA 1369
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY GGAGCTTTGATGCTCCACACTTCAACTCAAAGGCCATGTTATCGTCGTCGTCACAAA 1429
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY GGAAGTGGAAACCTTGAACTCGTGCTGTAAGAAAGAGCAACAACAGAGGGACCGCG 1489
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY GAAGAAGAGGAGCAGCAAGACGAAGAGAGAGGAGGAGTAACAGAGAGGTGCGTAGGTAC 1549
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
481 GluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValAlaArgGlyr 500
QY ACAGCGAGGTGAAGGAAGCGATGTGTTTCATCATGCCAGCTCATCCAGTAGGCCATC 1609
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaHisProValAlaIle 520
QY AAGCTTCTCCGAACTCCATCTGCTGGCTGGTATCAACCGTGAACCAACACACACAGA 1669
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY ATCTTCTTCAGGTGATAAGGCAATGTGATAGACCAAGATAGAGAGCAAGCGAAGGAT 1729
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY TTAGCATTCCTGGGTGGGTGAACAAGTTTGAGAGGCTCATCAAAAAACCAAGAGGAATCT 1789
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY CACTTTGTAGTGTCTGCTCTCAATCTCAATCTCGTCTGCTCTGAGAAAGAG 1849
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
581 HisPheValSerAlaArgProGlnSerGlnSerProSerProSerProGluLysGlu 600
QY TCTCCTCAGAAAGAGGATCAAGAGGAGGAAAAACCAAGGAGGAGGTCCACTCTTCA 1909

Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyGlyProLeuLeuSer 620
QY 1910 ATTTGAAGCTTTTAAC 1927
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
621 IleLeuLysAlaPheAsn 626
RESULT 9
ABU52571
ID ABU52571 standard; protein; 626 AA.
XX ABU52571;
XX 10-MAR-2003 (first entry)
XX Peanut Ara h1 mutant R91A.
XX Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;
KW mutein; anaphylactic food allergen; antiallergenic; vaccine;
KW wound healing.
XX Homo sapiens.
OS XX
PN WO200274250-A2.
XX 26-SEP-2002.
XX 18-MAR-2002; 2002WO-US009108.
XX 16-MAR-2001; 2001US-0276822P.
XX 18-MAR-2002; 2002US-00276822.
XX (PANA-) PANACEA PHARM.
XX Caplan M, Sosin H, Sampson H, Bannon G, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabjohn PA, Shin DS, Stanley JS;
XX WPI; 2003-018765/01.
XX
PT New modified anaphylactic food allergen, useful for preventing or
PT treating allergic reactions associated with e.g. anaphylactic allergens.
XX
XX Example 5; Page; 300pp; English.
XX
CC The invention relates to a modified anaphylactic food allergen has an
CC amino acid sequence that is substantially identical to that of natural
CC anaphylactic food allergen, except for a cysteine residue that has been
CC modified so that it cannot participate in the disulphide bond. The
CC modification may also comprise mutation of the IgE binding sites to
CC reduce allergenicity. Also included are: (1) a method of making a
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
CC or for causing a site specific mutation in the modified anaphylactic food
CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified
CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a food
CC allergen, mutated to alter its IgE binding characteristics. Note: The
CC present sequence is not shown in the specification but was created by the
CC indexer using information provided in the specification
XX
SQ Sequence 626 AA;

Alignment Scores: 5.96e-304 Length: 626
Pred. No.: 3280.00 Matches: 625
Score: 99.8% Conservative: 0
Percent Similarity: 99.8% Mismatches: 1
Best Local Similarity: 99.8%

Query Match: 91.5% Indels: 0
DB: 6 Gaps: 0
US-10-728-323-1 (1-2032) x ABUS2571 (1-626)

QY 50 ATGAGAGGAGGTTTCTCCACTGATGCTGTCTCTAGGATCCTTCTCTGGCTTCAGTT 109
DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyLeuValLeuAlaSerVal 20

QY 110 TCTGCACCGCATCCCAAGTATCATCACCTTACCAGAAAGAAAGAGAGAACCCCTCGCCAG 169
DB 21 SerAlaThrHisAlaIysSerProTyrGlnLysThrGluAsnProCysAlaGln 40

QY 170 AGTGTCTCCAGAGTTGTCAACAGAAACCGGATGACTTGAAGCAAAAGGCATCGAGTCT 229
DB 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60

QY 230 CGCTGCACCAAGCTCAGATATGATCCTCTGTGTCTATGATCCTCAGAGACACATGGC 289
DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80

QY 290 ACCACCAACCAACGTTCCCTCCAGGGGAGCGACACGTCGGCCCAACCCGGAGACTAC 349
DB 81 ThrThrAsnGlnArgSerProGlyGluAlaThrArgGlyArgGlnProGlyAspTyr 100

QY 350 GATGATCACCGCTCAACCCCGAAGAGAGAGGAGCGCGATGGGGACAGCTGGACCG 409
DB 101 AspAspAspArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGlyPro 120

QY 410 AGGAGCGGTAAAGAGAAGAAGACTGGAGAACCAAGAGAAGATTGGAGGGACCAAGT 469
DB 121 ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgProSer 140

QY 470 CATCACACCCCGAAGATTAAGCCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160

QY 530 CACAGTAGGCATGTAGGGAAGAAACATCTCGGAACAAACCTTCTACTTCCCGTCAAGG 589
DB 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180

QY 590 CGGTTTACGACCCGCTACGGGAACCAAAACGGTAGGATCCGGTCTCTGAGAGTTTGAC 649
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200

QY 650 CAAAGCTCAGGCAGTTTCAGATCTCCAGATCCACCGTATTGTGAGATCCGAGGCCAAA 709
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220

QY 710 CCTAACACTCTTCTTCTCCAGCACGCTGATGCTGATAACATCTCTGTATCCAGCAA 769
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGln 240

QY 770 GGGCAAGCACCGTACCGTAGCAATGGCAATACAGAAAGAGCTTTAATCTTGACGAG 829
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260

QY 830 GGCATCGCATCAGATCCATCCGCTTCTCATTTCTTCTACATCTTGAAACCGCATGACA 889
DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280

QY 890 CAGAACCTCAGAGTAGTAAAATCTCCATGCCCGTTAAACACACCCCGCCAGTTTGAAGAT 949
DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300

QY 950 TTTCTCCCGGAGCAGCGAGACCAATCATCTACTTCAGGGCTTCAGAGGAATACG 1009
DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320

QY 1010 TTGAGCGCCCTTCAATCGGAAATTCATGATACGATACGAGGGGTGCTGTAGAAGAAAT 1069
DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340

QY 1070 GCAGGAGGTGAGCAAGAGAGAGAGAGCGGAGCGGATCGAGTACTCGGAGTAGTGAGAAC 1129
DB 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360

QY 1130 AATGAGGAGTGATAGTCAAAAGTGTCAAAGAGACGTTTGAAGAACTTACTAAGCAGCT 1189
DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380

QY 1190 AATTCGCTCTCAAGAAAGGCTCCGAAGAGAGGAGATATCACCAACCAATCAACTTG 1249
DB 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400

QY 1250 AGAGAGCGGAGCCCGATCTTTTAAACAACCTTTGGGAAGTTATTTGAGGTGAAGCCAGAC 1309
DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420

QY 1310 AAGAAGAACCCCGAGCTTCAGACCTGGACATGCTCACCTGTGTAGAGATCAAGAA 1369
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440

QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGGCCATGTTTATCGTCTCGTCAACAAA 1429
DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460

QY 1430 GGAACCTGGAACCTTTGAACCTCGTGTGTAAAGAAAGAGCAACAACAGAGGGGACCGCG 1489
DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480

QY 1490 GAAGAAGAGGAGGACGAAGACCAAGAGAGGAGGAAGTTAACAGAGAGGTGGTAGGTAC 1549
DB 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500

QY 1550 ACAGCGAGGTTCAAGGAGGCGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGGCATC 1609
DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaHisProValAlaIle 520

QY 1610 AACGCTTCTCCGAACTCCATCTGCTGTGGCTTCGTTATCAACGCTGAAACCAACACAGA 1669
DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540

QY 1670 ATCTTCTTGCAGGTGATAAGACAAATGTGATAGACCATAGACATAGAGAACGGAAGGAT 1729
DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560

QY 1730 TTAGCATTCCTCGGTTCGGTGCAACAAAGTTGAGAAGCTCATCAAAACCAAGAGGAATCT 1789
DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580

QY 1790 CACTTTGTGAGTCTCGTCTCTCAATCTCAATCTCAATCTCGTCTCTCGTCTCGAAGAGAG 1849
DB 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600

QY 1850 TCTCTGAGAAAGAGATCAAGAGAGGAGGAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
DB 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyLysGlyProLeuLeuSer 620

QY 1910 ATTTTGAAGGCTTTTAAAC 1927
DB 621 IleLeuLysAlaPheAsn 626

RESULT 10
ABUS2574
ID ABUS2574 standard; protein; 626 AA.
XX AC ABUS2574;
XX DT 10-MAR-2003 (first entry)
XX DE Peanut Ara h1 mutant R499A.
XX KW Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;
KW mutain; anaphylactic food allergen; antiallergenic; vaccine;
KW wound healing.

XX OS Homo sapiens.
XX PN WO200274250-A2.
XX PD 26-SEP-2002.
XX PF 18-MAR-2002; 2002WO-US0009108.
XX PR 16-MAR-2001; 2001US-0276822P.
XX PR 18-MAR-2002; 2002US-00276822.
XX PA (PANA-) PANACEA PHARM.
XX PI Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
XX PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
XX PI Rabjohn PA, Shin DS, Stanley JS;
XX XX
XX WI; 2003-018765/01.
XX
XX New modified anaphylactic food allergen, useful for preventing or
XX treating allergic reactions associated with e.g. anaphylactic allergens.
XX
XX Example 5; Page; 300pp; English.
XX
XX The invention relates to a modified anaphylactic food allergen has an
XX amino acid sequence that is substantially identical to that of natural
XX anaphylactic food allergen, except for a cysteine residue that has been
XX modified so that it cannot participate in the disulphide bond. The
XX modification may also comprise mutation of the IgE binding sites to
XX reduce allergenicity. Also included are: (1) a method of making a
XX modified anaphylactic food allergen; (2) a nucleotide molecule encoding
XX or for causing a site specific mutation in the modified anaphylactic food
XX allergen; (3) a transgenic plant or animal expressing the modified
XX anaphylactic food allergen; (4) a method of treating an individual by
XX reducing the clinical response to a natural anaphylactic food allergen;
XX and an isolated fragment of peanut allergen Ara h 1. The modified
XX anaphylactic food allergen is useful for preventing or treating allergic
XX reactions associated with any natural allergen such as food, insect,
XX rubber or preferably anaphylactic allergens. It is also useful for
XX treating wounds in mammals such as bovine, canine, feline, caprine,
XX ovine, porcine, murine or equine species. The present sequence is a food
XX allergen, mutated to alter its IgE binding characteristics. Note: The
XX present sequence is not shown in the specification but was created by the
XX indexer using information provided in the specification
XX
XX Sequence 626 AA;

Alignment Scores:
Pred. No.: 5.96e-304 Length: 626
Score: 3280.00 Matches: 625
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 91.5% Indels: 0
DB: 6 Gaps: 0

US-10-728-323-1 (1-2032) x ABU52574 (1-626)
QY 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTGCTAGGGATCCCTGTGCTCGGTTTCAGTT 109
DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyLeuValLeuAlaSerVal 20
QY 110 TGTGCAACGGATGCAAGTCAATCCCTTACAGAGAAACAGAGAACCCCTGCGCCAG 169
DB 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGGTGCTCCAGAGTTGTCACAGAACCGGATGACTTTGAAGCAAAAGGATCGAGTCT 229
DB 41 ArgCysLeuGlnSerCysGlnGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGTGTCAACAGCTCGAGTAGTATGATCCCTGTTGTCTATGATCCTCGAGACACACTGGC 289
DB 61 ArgCysThrLysLeuGluTyrAspProArgProCysValTyrAspProArgGlyHisThrGly 80

QY 290 ACCACCAACAAGTTCCCTCCAGGGAGCGGACACACGTCGGCCCAACCCGGAGACTAC 349
DB 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCGTCAACCCCGAAGAGAGAGGAGGCGCGATGGGAGCAGCTGACCG 409
DB 101 AspAspArgArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGlyPro 120
QY 410 AGGAGCGTGAAGAGAAAGACTGGAGACAAACCAAGAGAAAGATTGGAGGCGACCAAGT 469
DB 121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
QY 470 CATCAGCAGCCCGAANAATAAGCCCGAAGAGAGAGAGAGAAACAAGAGTGGGACACA 529
DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
QY 530 CCAGGTAGCATGTGAGGAGAAACATCTCGGAACAACCTTTCTACTTCCCGTCAAGG 589
DB 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCGCTACCGGAAACCAAAACGGTAGGATCCGGTCTCTGAGAGGTTTAC 649
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGTCAAGCAGTTTCAGAAATCTCCAGAAATCACCGTATTCTGCAGATCGAGGCCAAA 709
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTGTTCTTCCCAAGCACGCTGATGCTGATAACAATCTTGTATCCAGCAA 769
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTAGCCCTAGCAATGCGAATAACAGAAAGAGCTTTAATCTTGACGAG 829
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCCTGCACTCAGATCCATCCCGTTTCATTTCTTACATCTTGAACCGCCATGACAAC 889
DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACTCAGAGTACTAAAATCTCCATGCCCGTTAAACACACCCGCCAGTTTGAGGAT 949
DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCGAGAGCCGAGACCAATCATCTTCTGAGGGCTTTCAGCAGGAATACG 1009
DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGCGCCCTTCAATCGGNAATTCAATGAGATACGGAGGCTGTGTTAGAGAGAAT 1069
DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 340
QY 1070 GCAGGAGGTGAGCAAGAGAGAGAGCGGAGCGAGCTGAGTAGTACTCGGAGTAGTGAAC 1129
DB 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGGAGTAGTCAAGTGTCAAAGGAGCAGCTTGAAGAACTTACTAGACGCT 1189
DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCGTCTCAAGAAAGCTCCGAGAGAGAGGAGATATCACCAACCCCAATCAACTTG 1249
DB 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGAGGCGAGCCCGATCTTTCTAAACAACCTTTGGGAAGTATTTCAGGTGAAGCCAGAC 1309
DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAACCCCGAGCTTCAGGACCTGGACATGCTGCTACCTGCTGTAGAGATCAAGAA 1369
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440

QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAAGCCCATGTTATCGTCGTCAACAAA 1429
Db GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAAGCTGGAACCTTGAACCTCGGTGCTGAAGAAAAGAGCAACACAGAGGGACGGCG 1489
Db GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAAGAAGAGGAGGACCAAGCAAGCAAGAGAGGAGGGAAGTAAACAGAGAGGTGCGTAGGTAC 1549
Db GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgAlaTyr 500
QY 1550 ACAGCAGGTGTAAGAAAGCGATGTGTTTCATCATCCAGCCAGCTCATCCAGTAGCCATC 1609
Db ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AAGCTTCTCCGAACCTCATCTGCTTGGCTTCGGTATCAACGCTGAAAAACACACAGA 1669
Db AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTCAGCTGATGAAGCAATGTGATAGACAGATAGAGAAGCAAGCGAAGGAT 1729
Db IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGTGCGGTGAACAAAGTTGAGAAGCTCATCAAAAACCAAGAGGAATCT 1789
Db LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTGCTCGTCTCAATCTCAATCTCAATCTCCGTCGTCCTGAGAAGAG 1849
Db HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCCTGAGAAAGAGGATCAAGAGGAGGAGAAACCAAGGAGGAGGAGGTCCTCCTTTCA 1909
Db SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGCTTTTAAAC 1927
Db IleLeuLysAlaPheAsn 626
RESULT 11
ABU52573
ID ABU52573 standard; protein; 626 AA.
AC ABU52573;
XX
DT 10-MAR-2003 (first entry)
XX Peanut Ara h1 mutant R109A.
DE Homo sapiens.
KW Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;
KW mutein; anaphylactic food allergen; antiallergenic; vaccine;
KW wound healing.
XX
OS
XX
PN WO200274250-A2.
XX
PD 26-SEP-2002.
XX
PF 18-MAR-2002; 2002WO-US0009108.
XX
PR 16-MAR-2001; 2001US-0276822P.
PR 18-MAR-2002; 2002US-00276822.
XX
XX (PANA-) PANACEA PHARM.
XX Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabin PA, Shin DS, Stanley JS;
XX WPI; 2003-018765/01.
DR
XX

PT New modified anaphylactic food allergen, useful for preventing or
PT treating allergic reactions associated with e.g. anaphylactic allergens.
XX Example 5; Page; 300pp; English.
XX
CC The invention relates to a modified anaphylactic food allergen has an
CC amino acid sequence that is substantially identical to that of natural
CC anaphylactic food allergen, except for a cysteine residue that has been
CC modified so that it cannot participate in the disulphide bond. The
CC modification may also comprise mutation of the IgE binding sites to
CC reduce allergenicity. Also included are: (1) a method of making a
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
CC or for causing a site specific mutation in the modified anaphylactic food
CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified
CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a food
CC allergen, mutated to alter its IgE binding characteristics. Note: The
CC present sequence is not shown in the specification but was created by the
CC indexer using information provided in the specification
XX
SQ Sequence 626 AA;
Alignment Scores:
Pred. No.: 5,96e-304 Length: 626
Score: 3280.00 Matches: 625
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 91.5% Indels: 0
DB: 6 Gaps: 0
US-10-728-323-1 (1-2032) x ABU52573 (1-626)
QY 50 ATGAGAGGGAGGGTTTCTCCACTGATGCTGTGTAGGATCTTGTCTCTGGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTCAACGCGATGCCAAGTCATCACCCTTACCAGAGAAACAGAGAACCCCTCGGCCAG 169
Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGGTGCTCCAGAGTTGTCAACAGSAACCGATGACTTGAACAAAAGGCGATCGAGTCT 229
Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGCTGCACCAAGCTCGAGTATGATCTCGTTCTGTCTATGATCTCGAGGACACACTGGC 289
Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAACGTTCCCTCCAGGGAGCGGACACGTCGGCCGCCAACCCGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATACCGCGCTCAACCCCGAAGAGAGGAGGAGGCGCGATGGGACAGCTGACCG 409
Db 101 AspAspAspArgArgGlnProArgAlaGluGluGlyGlyArgTyrpGlyProAlaGlyPro 120
QY 410 AGGGAGCGTGAAG 469
Db 121 ArgGluArgGluArgGluArgGluArgGluArgGluArgGluArgGluArgGluArgGlu 140
QY 470 CATCAGCAGCCACGGAAAAATAAGGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGlnGluThr 160
QY 530 CCAGTAGCATGTGAGGAG 589
Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180

QY 590 CGGTTTAGCACCCGCTACCGGAACCAAAACGGTAGATCCGGTCTCGCAGAGGTTTGAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAGGTCAGGCGAGTTTCAGAAATCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAAACACTCTGTCTCTCCCAAGCAGCGCTGATGCTGATAACATCTCTGTTATCCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTAGCCGAGCAAAATGGCAATACAGAAAGAGCTTTAATCTTGACGAG 829
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCATGCACTCAGAACTCCATCCGTTTTCATTTCTTACATCTTGACCGCCATGACAAC 889
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACTCAGAGTAGTAAATCTCCATGCCCGTTAAACACACCCCGCCAGTTTGAGGAT 949
Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTTCTCCCGCGAGCAGCCGAGACCAATCATCTTGTGAGGCTTCAGCAGGAATACG 1009
Db 301 PhePheProAlaSerSerArgGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGCCCGCTCAATGCGGATTCATGAGTACGAGGCTGCTGTAGAGAGAAT 1069
Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGGTGAGCAAGAGGAGAGGCGAGCGATGGAGTACTCGAGTAGTGAGAAC 1129
Db 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGGAGTGATAGTCAAAAGTGTCAAGAGACGACGCTTGAAGAACTTACTAGACGCT 1189
Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCGCTCAAGAGAGGCTCGAGAGAGGAGAGATATCACCAACCCCACTCACTTG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAGAGCGGAGCCCGATCTTTCTAAACAACCTTTGGGAAGTTATTTGAGGTGAAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAGAACCCCGCTTCAGGACCTGGACATGATGCTCACTGCTGATAGATCAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCGTCGTCGCAACAAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValAlaValAsnLys 460
QY 1430 GGAACTCGAAACCTTGAACCTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGACCGCG 1489
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAAGAAGAGGAGCGAAGCAAGAGAGAGGGAAGTACAGAGAGGTCGTAGGTAC 1549
Db 481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValAlaArgTyr 500
QY 1550 ACAGGAGGTTGAAGGAGGCGATGTTTATCATATGTCAGCAGCTCATCAGTAGCCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGAACCTCATCTGCTTGGCTTCGGTATCAACGCTGAAAAACAACACAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540

QY 1670 ATCTTCCTTCAGGTGATGAAGGACAAATGTGATAGACCAGATAGAGCAAGCGAAGGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCAATTCCTGGGTGCGGTCAACAAAGTTGAGAAGCTCATCAAAAACCAAGAGGAATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTCTGCTCCTCAATCTCAATCTCAATCTCCTGCTGCTCTGAGAAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCCTGCAAAAGAGGATCAAGGAGGAAACCAAGGAGGAGGAGGAGGTCCTCCTTTCA 1909
Db 601 SerProGluLysGluAspGlnGluGluGlnGlnGlnGlyGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTTAAAC 1927
Db 621 IleLeuLysAlaPheAsn 626
RESULT 12
ADM12053
ID ADM12053 standard; protein; 626 AA.
XX
AC ADM12053;
XX
XX 20-MAY-2004 (first entry)
DT
XX Arachis hypogaea 1 (Ara h1) protein.
DE
XX antigen presentation enhancing hybrid polypeptide; mammalian Ii-Key;
KW MHC class II; antibacterial; virucide; fungicide; antirheumatic;
KW antiarthritic; neuroprotective; dermatological; immunosuppressive;
KW antiinflammatory; antidiabetic; antithyroid; immune;
KW rheumatoid arthritis; multiple sclerosis; lupus erythematosus;
KW diabetes mellitus; myasthenia gravis; autoimmune thyroiditis;
KW scleroderma; dermatomyositis; pemphigus.
XX
XX Arachis hypogaea.
OS
XX US2003235594-A1.
PN
XX
PD 25-DEC-2003.
XX
PF 17-SEP-2002; 2002US-00245871.
XX
PR 14-SEP-1999; 99US-00396813.
PR 17-JUL-2002; 2002US-00197000.
XX
PA (ANTI-) ANTIGEN EXPRESS INC.
XX
XX Humphreys R, Xu M;
XX WPI; 2004-070554/07.
DR
XX Novel Ii-Key/antigen presentation enhancing hybrid polypeptide, useful
PT for treating infections, rheumatoid arthritis, multiple sclerosis, lupus
PT erythematosus and diabetes mellitus.
XX
XX Example 1; Page 16-17; 87pp; English.
XX
CC The invention relates to a novel antigen presentation enhancing hybrid
CC polypeptide. The novel polypeptide has an N-terminal element consisting
CC of 4-16 residues of a mammalian Ii-Key peptide and its non-N-terminal
CC deletion modifications, a chemical structure covalently linking the N-
CC terminal element to an MHC class II-presented epitope of a C-terminal
CC element. The C-terminal element comprises an antigenic epitope, which
CC binds to an antigenic peptide binding site of an MHC class II molecule.
CC The antigen presentation enhancing hybrid polypeptide has the following
CC activities: antibacterial, virucide, fungicide, antirheumatic,
CC antiarthritic, neuroprotective, dermatological, immunosuppressive,
CC antiinflammatory, antidiabetic, and antithyroid. The antigen presentation
CC enhancing hybrid polypeptide is useful for modulating the immune response

CC in an individual and for treating infections (such as bacteria, virus,
 CC parasite and fungus), rheumatoid arthritis, multiple sclerosis, lupus
 CC erythematous, diabetes mellitus, myasthenia gravis, autoimmune
 CC thyroiditis, scleroderma, dermatomyositis and pemphigus. This sequence
 CC represents a mammalian II key related protein of the invention.

XX SQ Sequence 626 AA;

Alignment Scores:
 Pred. No.: 7 43c-304 Length: 626
 Score: 3279.00 Matches: 625
 Percent Similarity: 99.8% Conservatives: 0
 Best Local Similarity: 99.8% Mismatches: 1
 Query Match: 91.4% Indels: 0
 DB: 8 Gaps: 0

US-10-728-323-1 (1-2032) x ADM12053 (1-626)

QY 50 ATGAGAGGAGGTTCTCCACTGATGCTTCTAGGATGCTTCTCGGCTTCAGTT 109
 DB 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyLeuValLeuAlaSerVal 20
 QY 110 TCTGCCACCATGCCAAGTATCATCCTTACCAGAAAGAACAGAGAACCCCTCGCCAG 169
 DB 21 SerAlaThrHisAlaLysSerProTyrGlnLysThrGluAsnProCysAlaGln 40
 QY 170 AGGTGCTCCAGAGTTGTCAACAGGAACCGATGACTTGAAGCAAAAGGCATCGAGTCT 229
 DB 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlyLysAlaCysGluSer 60
 QY 230 CGTGCACCAAGTCCAGTATGATCTCTGTTGTCTATGATCCTCGAGGACACACTGGC 289
 DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
 QY 290 ACCACCAACCAAGTTCCTCCAGGGAGCGACAGCTGGCCGCCAACCCGAGACTAC 349
 DB 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
 QY 350 GATGATGACCGCGTCAACCCGAGAGAGAGGAGCGCGATGGGACCGAGCTGGACCG 409
 DB 101 AspAspAspArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGlyPro 120
 QY 410 AGGAGCGTGAAGAGAAGAACTGGAGACAAACCAAGAGAAGATTGGAGGCGACCAAGT 469
 DB 121 ArgGluArgGluArgGluArgTrpArgGlnProArgGluAspTrpArgProSer 140
 QY 470 CATCAGCAGCCAGGAATAAGCCCGAAGAGAGAGAGAGAACAAAGATGGGAAACA 529
 DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
 QY 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAAACCTTCTACTTCCCGTCAAGG 589
 DB 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
 QY 590 CGGTTTACGACCCGCTACGGGAACCAAAACGGTAGGATCGGGTCTCGAGAGGTTTGAC 649
 DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
 QY 650 CAAGGTCAAGGCAGTTTCAGATCTCCAGATCACCCTATGTGTGAGATCGAGGCCAAA 709
 DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
 QY 710 CCTAACACATCTTCTTCCCAAGCAGCTGATGCTGATAACATCTTGTATCCAGCAA 769
 DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnLeuValIleGlnGln 240
 QY 770 GGGCAAGCCACCGTAGCCGTAGCAATGGCAATACAGAAAGAGCTTTAATCTTGACGAG 829
 DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
 QY 830 GGCATGCACTCAGATCCCATCCGGTTTCATTTCTTACATCTTGAACCCCATGACAAC 889
 DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280

QY 890 CAGAACCTCAGAGTAGCTAAAAATCTCCATGCCCGTTAAACACACCCGCGCAGTTGAGGAT 949
 DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
 QY 950 TTCTCCCGGCGAGCAGCGAGACCAATCATCTTCTGCGGGCTTCAGCAGGAATACG 1009
 DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
 QY 1010 TTGGAGCGCGCTTCAATCGCGAATTCATATGATACGAGGAGTCTGTTAGAGAGAAT 1069
 DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 340
 QY 1070 GCAGGAGGTGAGCAAGAGAGAGAGCGGAGCGCATGATGAGTACTCGGAGTAGTGAGAAC 1129
 DB 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgTrpSerThrArgSerSerGluAsn 360
 QY 1130 AATGAAGAGTAGTATGATCAAACTGTCAAAGAGACAGCTTGAAGAACTTACTAAGCAGCT 1189
 DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
 QY 1190 AATCCGTCTCAAGAAAGGCTCCGAGAGAGAGGAGATATCACCAACCAATCAACTTG 1249
 DB 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
 QY 1250 AGAGAGCGGAGCGCGCATCTTTCTAACTTTTGGGAAGTATTGAGGTGAAGCCAGAC 1309
 DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
 QY 1310 AAGAAGAACCCCGACTTCAGACCTCGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369
 DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu 440
 QY 1370 CGAGCTTTGATGCTCCACACTTCAACTCAAGAGCCATGGTTATCGTCTCGTCAACAA 1429
 DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
 QY 1430 GGAACTTGGAAACTCTGACTCGTGTAAAGAAAGAGACAAACAGAGGGGACCGCGG 1489
 DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
 QY 1490 GAAGAGAGGAGGAGCAAGACGAAAGAGAGGAGGAGTAACAGAGAGGTGCTGAGTAC 1549
 DB 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
 QY 1550 ACAGCGAGGTTCAGAGGCGCATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGGCATC 1609
 DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
 QY 1610 AACGCTTCTCCGAACTCCATCTGCTTGGCTTTCGGTATCAACGCTGAAACCAACACAGA 1669
 DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
 QY 1670 ATCTCTCTTGCAGGTGATAAGACAAATGTGATAGACCATAGATAGAGAACGAGGAAGGAT 1729
 DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
 QY 1730 TTAGCATTTCCCGGTCCGGTCAACAGTTGAGAGCTCATCAAAAACAGAGGAAGTCT 1789
 DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
 QY 1790 CACTTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCCGCTTCTCCGCTGAGGAAGAG 1849
 DB 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
 QY 1850 TCTCTGAGAAAGAGATCAAGAGAGGAGAAACCAAGAGGAGGAGGTCCTCTCTTTCA 1909
 DB 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
 QY 1910 ATTTTGAAGCTTTTAAAC 1927
 DB 621 IleLeuLysAlaPheAsn 626

RESULT 13	ABU52569	ABU52569 standard; protein; 626 AA.
XX	AC	ABU52569;
XX	DT	10-MAR-2003 (first entry)
XX	DE	Peanut Ara h1 mutant D52A.
XX	KW	Allergy; Ara h1; Ara h2; Ara h3; IGE binding site; peanut; mutant; munein; anaphylactic food allergen; antiallergenic; vaccine; wound healing.
XX	OS	Homo sapiens.
XX	PN	W0200274250-A2.
XX	PD	26-SEP-2002.
XX	PF	18-MAR-2002; 2002WO-US009108.
XX	PR	16-MAR-2001; 2001US-0276822P.
XX	PR	18-MAR-2002; 2002US-00276822.
XX	PA	(PANA-) PANACEA PHARM.
XX	PI	Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G; Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki S; Rabjohn PA, Shin DS, Stanley JS;
XX	PI	WPI; 2003-018765/01.
XX	DR	New modified anaphylactic food allergen, useful for preventing or treating allergic reactions associated with e.g. anaphylactic allergens.
XX	PS	Example 5; Page; 300pp; English.
XX	CC	The invention relates to a modified anaphylactic food allergen has an amino acid sequence that is substantially identical to that of natural anaphylactic food allergen, except for a cysteine residue that has been modified so that it cannot participate in the disulphide bond. The modification may also comprise mutation of the IGE binding sites to reduce allergenicity. Also included are: (1) a method of making a modified anaphylactic food allergen; (2) a nucleotide molecule encoding or for causing a site specific mutation in the modified anaphylactic food allergen; (3) a transgenic plant or animal expressing the modified anaphylactic food allergen; (4) a method of treating an individual by reducing the clinical response to a natural anaphylactic food allergen; and an isolated fragment of peanut allergen Ara h 1. The modified anaphylactic food allergen is useful for preventing or treating allergic reactions associated with any natural allergen such as food, insect, rubber or preferably anaphylactic allergens. It is also useful for treating wounds in mammals such as bovine, canine, feline, caprine, ovine, porcine, murine or equine species. The present sequence is a food allergen, mutated to alter its IGE binding characteristics. Note: The present sequence is not shown in the specification but was created by the indexer using information provided in the specification
XX	SQ	Sequence 626 AA;
XX	Alignment Scores:	
XX	Pred. No.:	9.26e-304
XX	Score:	3278.00
XX	Percent Similarity:	99.8%
XX	Best Local Similarity:	99.8%
XX	Query Match:	91.4%
XX	DB:	6
XX	US-10-728-323-1 (1-2032) x ABU52569 (1-626)	
XX	50 ATGAGAGGAGGGTTTCTCCACTGATCTGTGCTAGGGATCCTTGCTGGCTTCAGTT	109

```
QY 1190 AATCCGCTCTCAAGAAAGGCTCCGAAGAAGAGGAGATATCACCAACCAATCAACTTG 1249
DB 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGAGGAGCCCGATCTTCTTAAACAATTTGGGAAGTTATTTGAGTGAGGCAGAC 1309
DB 401 ArgGluGluProAspLeuSerAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsn 420
QY 1310 AAGAAGAACCCAGCTTCAGACCTCGGACATGCTCACCTGCTAGAGATCAAGAA 1369
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCTGCTCGTCAACAAA 1429
DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAATCGAAACCTTGAACTCGTGGCTGTAAGAAAAGACAAACAGAGGGGAGCGGG 1489
DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAAGAGAGGAGACCAAGACGAGAGAGAGGAGGAGTAACAGAGAGTGGTAGGTAC 1549
DB 481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgGly 500
QY 1550 ACAGCGAGGTGAAGGAAGCGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCATC 1609
DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGTTCTTCGAATCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1669
DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTCAGGTGATAGGACATCTGATAGACCATAGACCATAGAGCAAGCGAGGT 1729
DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGTGCGGTGAACAAAGTTGAGAAGCTCATCAAAAACAGAGGAATCT 1789
DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTGTCTGCTCTCAATCTCAATCTCAATCTCGTCTCGTCTGAGAAAGAG 1849
DB 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTGAGAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
DB 601 SerProGluLysGluAspGlnGluGluGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGCTTTTAAAC 1927
DB 621 IleLeuLysAlaPheAsn 626
RESULT 14
ID ABUS2572
XX standard; protein; 626 AA.
AC ABUS2572;
XX
XX 10-MAR-2003 (first entry)
XX Peanut Ara h1 mutant D103A.
XX
XX Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;
XX mutein; anaphylactic food allergen; anti-allergenic; vaccine;
XX wound healing.
XX
XX Homo sapiens.
XX
XX WO200274250-A2.
XX
XX 26-SEP-2002.
XX
```

```
PF 18-MAR-2002; 2002WO-US009108.
XX
PR 16-MAR-2001; 2001US-0276822P.
PR 18-MAR-2002; 2002US-00276822.
XX
PA (PANA-) PANACEA PHARM.
XX
XX Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabjohn PA, Shin DS, Stanley JS;
XX
XX WPI; 2003-018765/01.
XX
PT New modified anaphylactic food allergen, useful for preventing or
PT treating allergic reactions associated with e.g. anaphylactic allergens.
XX
XX Example 5; Page: 300pp; English.
XX
XX The invention relates to a modified anaphylactic food allergen has an
XX amino acid sequence that is substantially identical to that of natural
XX anaphylactic food allergen, except for a cysteine residue that has been
XX modified so that it cannot participate in the disulphide bond. The
XX modification may also comprise mutation of the IgE binding sites to
XX reduce allergenicity. Also included are: (1) a method of making a
XX modified anaphylactic food allergen; (2) a nucleotide molecule encoding
XX or for causing a site specific mutation in the modified anaphylactic food
XX allergen; (3) a transgenic plant or animal expressing the modified
XX anaphylactic food allergen; (4) a method of treating an individual by
XX reducing the clinical response to a natural anaphylactic food allergen;
XX and an isolated fragment of peanut allergen Ara h 1. The modified
XX anaphylactic food allergen is useful for preventing or treating allergic
XX reactions associated with any natural allergen such as food, insect,
XX rubber or preferably anaphylactic allergens. It is also useful for
XX treating wounds in mammals such as bovine, canine, feline, caprine,
XX ovine, porcine, murine or equine species. The present sequence is a food
XX allergen, mutated to alter its IgE binding characteristics. Note: The
XX present sequence is not shown in the specification but was created by the
XX indexer using information provided in the specification
XX
XX Sequence 626 AA;
XX
Alignment Scores:
Pred. No.: 9,26e-304 Length: 626
Score: 3278.00 Matches: 625
Percent Similarity: 99.8% Conservativity: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 91.4% Indels: 0
DB: 6 Gaps: 0
US-10-728-323-1 (1-2032) x ABUS2572 (1-626)
QY 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTGCTAGGATCCTTGTCTCGCTTCAGTT 109
DB 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTCAACGCGATGCCAAGTCATCCTTACAGAGAAAGAAACAGAGAACCCCTCGCGCCAG 169
DB 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGGTGCTCCAGAGTTGTCAACAGAACCGGATGACTTGAAGCAAAAGCATCGCAGTCT 229
DB 41 ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGGC 289
DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCCAACCAACGCTTCCCTCCAGGAGCGGACACGTCGGCGCGCAACCCGAGACTAC 349
DB 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATACCGCGCTCAACCCCGAAGAGAGAGGAGCGCGATGGGACCACTGACCGC 409
```


Db 101 AspAspAlaArgArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGlyPro 120
QY 410 AGGACCGTGAAGAGAAGAACTGGAGACAAACCAAGAGAAGATTGGAGGCGCAACT 469
Db 121 ArgGluArgGluArgGluArgTrpArgGlnProArgGluAspTrpArgArgProSer 140
QY 470 CATCACAGCCACGGAAATAAGCCCGAAGGAGAGAGAGAACAAAGAGTGGGAACA 529
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGlnGluTrpGlyThr 160
QY 530 CCAAGTAGCCATGTAGGGAAGAAACATCTCGGAACAAACCTTTCACCTCCCGTCAAGG 589
Db 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCCGCTACGGGAACCAAAACCGTAGGATCCGGTCTCTGCAGAGTTTGAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CARAGGTCAAGGCAGTTTCAGATCTCCAGATCCCGTATTGTGCAGATCCGAGGCCAAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAAACACTCTGTTCTTCCCAAGCACGCTGATGCTGATAACATCTCTGTTATCCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGSCAAGCCACCGTAGCCGTAGCAATGGCAATPACAGAAAGAGCTTTAATCTTGACGAG 829
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCATGTCACTCAGATCCCATCCGTTTCATTTCTACATCTTGAACCGCCCATGACAAC 889
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACTCAGAGTAGTAAATCTCATGCCCGTTAAACACACCCGCCAGTTTGAGGAT 949
Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCAGCAGCCGAGACCAATCATCTACTTTCAGGGCTTCAGCGAATACG 1009
Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGAGGCCCGCTTCAATCGGAATTCATGAGATACGGGGTGTGTTAGAGAGAT 1069
Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 340
QY 1070 GCAGGAGGTGAGCAAGAGGAGAGGCGAGCGCATGGAGTACTCGAGTAGTGAGAAC 1129
Db 341 AlaGlyGlyGlnGlnGluArgGlyGlnArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGAGTAGTACAAAGTGTCAAGGAGCAGCTTGAAGAACTTACTAAGCGCGCT 1189
Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCGTCTCAAGAAAGCTCCGAGAGAGGAGATATCACCAACCAACTCACTTG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGCGGAGCCCGATCTTCTTAACTTTGGGAAGTTATTGAGGTGAAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAACCCCGAGTTCCAGGACCTGGACATGATGCTCACTGTGTGAGAGATCAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGCTTATCTCGTCTGTCACAAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAAGTGGAAACCTTGAACTCGTGGCTGTAAAGAAAAGAGCAACAAAGAGGGGACGCGG 1489
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480

QY 1490 GAAGAAGAGGAGGACCAAGACCAAGAGGAGGGAAGTAAACAGAGAGTGCGTAGGTAC 1549
Db 481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTTGAAGGAAGGCGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGAATCTCATCTGCTGGCTTCGGTATCAACGCTGAAACCAACACACAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCCTTCAGGTCATAAGACAACTGTGATAGACCATAGATAGACAGCAAGCGAAGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTTCCCTGGGTCCGGTGAAACAAGTTGAGAAAGCTCATCAAAACCCAGAAAGAAATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTGTCTCTCTCAATCTCAATCTCAATCTCCGTCTCTCCTGAGAAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCCTCAGAAAGAGATCAAGAGGAGGAAACCAAGGAGGAGGAGGTCCTCCTTTTCA 1909
Db 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTTAAAC 1927
Db 621 IleLeuLysAlaPheAsn 626
RESULT 15
AAW22150
ID AAW22150 standard; protein; 626 AA.
XX AC AAW22150;
XX DT 17-OCT-2003 (revised)
DT 29-DEC-1997 (first entry)
XX Peanut allergen Ara hi.
DE Peanut; seed storage protein; allergen; allergy; hypersensitivity;
KW vaccine; anaphylactic shock; immunotherapy; therapy; monoclonal antibody;
KW ELISA; analysis; Ara hi.
XX Arachis hypogaea; strain Florunner.
XX Key Location/Qualifiers
FH Peptide 1..22
FT /label= Sig_peptide
FT Protein 23..626
FT /label= Mat_protein
FT Modified-site 521..523
FT /note= "N-glycosylation site"
XX WO9724139-A1.
XX PN 10-JUL-1997.
XX PD 23-SEP-1996; 96WO-US015222.
XX PF 29-DEC-1995; 95US-0009455P.
XX PR 04-MAR-1996; 96US-00610424.
XX PA (UYAR-) UNIV ARKANSAS.
XX PI Burks AW, Helm RM, Cockrell G, Stanley JS, Bannon GA;
PI WPI; 1997-363453/33.
DR N-PSDB; AAT76613.
DR

XX Peanut allergens Ara hi and Ara hII - used for vaccination and in two-
PT site monoclonal antibody based ELISA.
XX Claim 31; Page 172; 354pp; English.
XX This polypeptide comprises major peanut allergen Ara hi (AAW22149). Its
CC sequence was deduced from cDNA clone P41b (AAT76613), isolated from
CC peanut seed cDNA using a primer (see AAT76616) based on an isolated Ara
CC hi peptide (see AAW24206). The sequence shows significant homology with
CC the vicilin family of seed storage proteins of other legumes. The
CC allergen is recognised by serum IgE from a large proportion of
CC individuals with peanut hypersensitivity. Ara hi and Ara hII (see
CC AAW24164) can be used to raise monoclonal antibodies which are used in a
CC specific two-site MAB ELISA for the detection of Ara hi or Ara hII
CC (claimed). IgE-binding Ara hi antigen epitopes (see AAW24165-87) may be
CC used in vaccines to protect against allergic reactions to peanut
CC allergens, e.g. anaphylactic shock. (Updated on 17-Oct-2003 to
CC standardise OS field)
XX
SQ Sequence 626 AA;

Alignment Scores:

Pred. No.: 2,23e-303 Length: 626
Score: 3274.00 Matches: 624
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 2
Query Match: 91.3% Indels: 0
DB: 2 Gaps: 0

US-10-728-323-1 (1-2032) x AAW22150 (1-626)

QY	50	ATCAGAGGAGGGTTCTCCACTGATGCTGTGTAGGATCCTTCTCTGGCTTCAGTT	109
DB	1	MetArgGlyArgValSerProLeuMetLeuLeuGlyLeuValLeuAlaSerVal	20
QY	110	TTGCAACGCATGCCAAGTCATCCTTACCAAGAAACACAGAACCCCTCGCCAC	169
DB	21	SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln	40
QY	170	AGTGCCTCCAGATGTTCAACAGNACCGGATGACTTGAAGCAAAAGCATCGAGTCT	229
DB	41	ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer	60
QY	230	CGTGCACCAAGCTCGAGTATGATCCTCGTGTGTCTATGATCCTCGAGGACACATGCG	289
DB	61	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly	80
QY	290	ACCACCAACCAAGTTCCTCCAGGGAGCGGACACGTTGGCCGCCAACCCCGGAGACTAC	349
DB	81	ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	100
QY	350	GATGATACCGCGCTCAACCCGAGAGAGGAGGCGCGATGGGACACAGCTGACCG	409
DB	101	AspAspAspArgArgGlnProArgGluGluGlyArgTyrGlyProAlaGlyPro	120
QY	410	AGGAGCGTGAAGAGAAAGACTCGAGACCAACCAAGAGAAGATTGGAGGCGACCACT	469
DB	121	ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgProSer	140
QY	470	CATCAGCACCGGAAATAAGGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	529
DB	141	HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGluGlnLutrpGlyThr	160
QY	530	CCAGGTAGCCATGTGAGGAAAGAAACATCTCGAACAACCCCTTCTACTCCCGTCAAGG	589
DB	161	ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg	180
QY	590	CGGTTTAGCACCGCTACCGGAACCAAAACGGTGGATCCGGTCTCTCGAGAGTTTGAC	649
DB	181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	200
QY	650	CAAAGGTCAAGGCAGTTTCCAGAAATCTCCAGAAATCACCCTATTGTGCAGATCGAGGCCAA	709

DB	201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys	220
QY	710	CCTAACACTCTGTCTTCCCAAGCACGCTGATGCTGATAACATCCTCTGCTTATCCAGCAA	769
DB	221	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	240
QY	770	GGGCAAGCCACCGTCACCGTACCAATGGCAATACAGAAAGAGCTTTAATCTTGACGAG	829
DB	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260
QY	830	GGCCATGCACCTCAGAAATCCATCCCGTTTCATTTCTTACATCTTGAACCCCATGACAAC	889
DB	261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	280
QY	890	CAGAACCCTCAGAGTAGCTAAATCTCCATGCCCGTTAAACACACCCCGCCAGTTGAGGAT	949
DB	281	GlnAsnLeuArgValAlaLysIleSerMetProValThrGlnProGlyGlnPheGluAsp	300
QY	950	TTCTTCCCGCGAGCGAGCCGAGACCAATCATCTTCTGACGGGCTTCAGCAGGAATACG	1009
DB	301	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	320
QY	1010	TTGAGAGCGCCTTCAATGCGGAATCAATGAGATACGAGGGTGTCTGTAGAGAGAAT	1069
DB	321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn	340
QY	1070	GCAGGAGGTGACAGAGGAGGAGGCGAGGCGGATCGGACTCGGAGTAGTGAGAAC	1129
DB	341	AlaGlyGlyGluGlnGluArgGlyGlnArgTyrPsrThrArgSerSerGluAsn	360
QY	1130	AATGAAGGAGTGATAGTCAAAAGTGTAAAGGAGCACGTTGAAAGAACTTACTAAGCACGT	1189
DB	361	AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	380
QY	1190	AAATCCGTCTCAAGAAAGGCTCCGAAGAGGAGAGATATACCAACCAATCAACTTG	1249
DB	381	LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu	400
QY	1250	AGAGAGGCGGCGCGATCTTCTAACCACTTTGGNAAGTATTATGAGGTGAGGCCAGAC	1309
DB	401	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	420
QY	1310	AAGAAAGAACCCCGACTTCAGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA	1369
DB	421	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu	440
QY	1370	GGAGCTTTGATGCTCCACACTTCAACTCAAGGCGCATGGTTATCGTCTGCTCAACAAA	1429
DB	441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys	460
QY	1430	GGAACTGGAAACCTTGAACCTCGTCTGTNAGAAAAGACACACAGAGGGGCGCGG	1489
DB	461	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg	480
QY	1490	GAAGAAGAGGAGGACCAAGACGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1549
DB	481	GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr	500
QY	1550	ACAGCGAGTTGAAGGAGGCGATGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC	1609
DB	501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520
QY	1610	AACGCTTCTCCGAACCTCCTGCTTGGCTTCGCTATCAACGCTGAAAACACACACAGA	1669
DB	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg	540
QY	1670	ATCTCTCTTGACGGTGATAAGGACAAATGTGTATAGACCAGATAGAGAGCAAGCAGGAT	1729
DB	541	IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp	560
QY	1730	TTAGCATTCCTCGGTGCGGTGAACAAGTTGAGAAGCTCATCAAAAACCAAGAGGAATCT	1789

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 22:18:32 ; Search time 12.2215 Seconds
(without alignments)
4123.811 Million cell updates/sec

Title: US-10-728-323-1

Perfect score: 3586

Sequence: 1 aataatcatatattatc.....cgttgtgcgttttcttc 2032

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p-model -DEV=xlp
-Q=/abs/ABSWEB_spool/US10728323/runat_15052006_172137_22463/app_query.fasta_1
-DB=issued_Patents_AA -QFMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs02p -USER=US10728323@cgn1.1.131@runat_15052006_172137_22463
-NCPU=6 -ICPU=3 -NO MAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA:*
1: /cgn2_6/prodata/1/iaa/5 COMB.pcp.*
2: /cgn2_6/prodata/1/iaa/6 COMB.pcp.*
3: /cgn2_6/prodata/1/iaa/H COMB.pcp.*
4: /cgn2_6/prodata/1/iaa/PTUS COMB.pcp.*
5: /cgn2_6/prodata/1/iaa/RE COMB.pcp.*
6: /cgn2_6/prodata/1/iaa/backfiles.pcp.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3278	91.4	626	2	US-09-106-872A-4
2	1700	47.4	335	2	US-09-106-872A-17
3	1279	35.7	605	1	US-07-955-905A-24
4	1267	35.3	571	1	US-07-955-905A-25
5	1171	32.7	449	2	US-10-245-227D-91
6	1169	32.6	449	2	US-10-245-227D-95
7	1169	32.6	449	2	US-10-245-227D-97
8	1163	32.4	439	2	US-10-245-227D-1
9	1163	32.4	449	2	US-10-245-227D-89
10	1162	32.4	449	2	US-10-245-227D-93
11	1157	32.3	417	2	US-10-245-227D-14
12	1157	32.3	425	2	US-10-245-227D-12

13	1117.5	31.2	390	2	US-10-245-227D-2	Sequence 2, Appli
14	1068	29.8	410	1	US-07-955-905A-26	Sequence 26, Appli
15	811	22.6	566	1	US-07-955-905A-2	Sequence 2, Appli
16	811	22.6	566	1	US-07-955-905A-27	Sequence 22, Appli
17	788	22.0	421	1	US-07-955-905A-22	Sequence 27, Appli
18	777	21.7	587	1	US-07-955-905A-23	Sequence 23, Appli
19	596.5	16.6	536	2	US-10-264-303-4	Sequence 4, Appli
20	595.5	16.6	540	2	US-10-264-303-3	Sequence 3, Appli
21	591.5	16.5	489	2	US-09-424-283-3	Sequence 1, Appli
22	569	15.9	524	2	US-09-424-283-1	Sequence 1, Appli
23	559.5	15.6	448	2	US-09-323-195A-18	Sequence 18, Appli
24	514.5	14.3	409	2	US-09-424-283-4	Sequence 4, Appli
25	513.5	14.3	523	2	US-09-323-195A-17	Sequence 17, Appli
26	501	14.0	444	2	US-09-424-283-2	Sequence 2, Appli
27	242.5	6.8	454	2	US-09-805-694B-4	Sequence 4, Appli
28	183	5.1	762	2	US-09-252-991A-29423	Sequence 29423, A
29	173	4.8	1564	2	US-10-144-198-2	Sequence 2, Appli
30	173	4.8	1564	2	US-10-144-198-4	Sequence 4, Appli
31	172	4.8	798	2	US-10-104-047-3335	Sequence 3335, Ap
32	166.5	4.6	720	2	US-09-252-991A-21881	Sequence 21881, A
33	164.5	4.6	1266	2	US-09-252-991A-30851	Sequence 30851, A
34	163	4.5	714	2	US-09-949-016-10296	Sequence 10296, A
35	160	4.5	630	2	US-09-252-991A-26324	Sequence 26324, A
36	160	4.5	801	2	US-09-252-991A-29274	Sequence 29274, A
37	160	4.5	1898	1	US-08-056-200-94	Sequence 94, Appli
38	160	4.5	1898	1	US-08-800-644-94	Sequence 1280, Ap
39	160	4.5	1898	2	US-09-538-092-1280	Sequence 1280, Ap
40	159	4.4	1043	2	US-09-252-991A-28885	Sequence 28885, A
41	158.5	4.4	1075	2	US-09-252-991A-18387	Sequence 18387, A
42	157	4.4	1780	1	US-08-769-309A-5	Sequence 5, Appli
43	157	4.4	1780	2	US-08-994-570-5	Sequence 5, Appli
44	157	4.4	1781	2	US-09-961-403-13	Sequence 13, Appli
45	157	4.4	2079	2	US-09-949-016-8301	Sequence 8301, Ap

ALIGNMENTS

RESULT 1
US-09-106-872A-4
; Sequence 4, Application US/09106872A
; Patent No. 6486311
; GENERAL INFORMATION:
; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Stanley, J. Steven
; APPLICANT: Cockrell, Gael
; APPLICANT: King, Nina E.
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Helm, Ricki M.
; APPLICANT: Bannan, Gary A.
; TITLE OF INVENTION: Peanut Allergens and Methods
; FILE REFERENCE: HS 103 CIP
; CURRENT APPLICATION NUMBER: US/09/106, 872A
; CURRENT FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: PCT/US96/15222
; PRIOR FILING DATE: 1996-09-23
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Arachis hypogaea
; FEATURE:
; OTHER INFORMATION: Amino Acids 25-34 are Ara H 1 binding epitope,
; OTHER INFORMATION: peptide 1
; OTHER INFORMATION: Amino Acids 48-57 are Ara H 1 binding epitope,
; OTHER INFORMATION: peptide 2
; OTHER INFORMATION: Amino Acids 65-74 are Ara H 1 binding epitope,
; OTHER INFORMATION: peptide 3
; OTHER INFORMATION: Amino Acids 89-98 are Ara H 1 binding epitope,
; OTHER INFORMATION: peptide 4
; OTHER INFORMATION: Amino Acids 97-106 are Ara H 1 binding epitope,
; OTHER INFORMATION: peptide 5
; OTHER INFORMATION: Amino Acids 107-116 are Ara H 1 binding epitope,

```
/ OTHER INFORMATION: peptide 6
/ OTHER INFORMATION: Amino Acids 123-132 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 7
/ OTHER INFORMATION: Amino Acids 134-143 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 8
/ OTHER INFORMATION: Amino Acids 143-152 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 9
/ OTHER INFORMATION: Amino Acids 294-303 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 10
/ OTHER INFORMATION: Amino Acids 311-320 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 11
/ OTHER INFORMATION: Amino Acids 325-334 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 12
/ OTHER INFORMATION: Amino Acids 344-353 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 13
/ OTHER INFORMATION: Amino Acids 393-402 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 14
/ OTHER INFORMATION: Amino Acids 409-418 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 15
/ OTHER INFORMATION: Amino Acids 461-470 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 16
/ OTHER INFORMATION: Amino Acids 498-507 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 17
/ OTHER INFORMATION: Amino Acids 525-534 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 18
/ OTHER INFORMATION: Amino acids 539-548 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 19
/ OTHER INFORMATION: Amino acids 551-560 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 20
/ OTHER INFORMATION: Amino acids 559-568 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 21
/ OTHER INFORMATION: Amino acids 578-587 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 22
/ OTHER INFORMATION: Amino acids 597-606 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 23
/ US-09-106-872A-4

Alignment Scores:
Pred. No.: 1,116-311 Length: 626
Score: 3278.00 Matches: 625
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 91.4% Indels: 0
DB: 2 Gaps: 0

US-10-728-323-1 (1-2032) x US-09-106-872A-4 (1-626)

QY 50 ATGAGAGGAGGGTTCTCCACTGATGCTGCTAGGATCCTGTCTCTGGCTTCAGTT 109
DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGCATGCCAAGTCATCACCTTACCAGAAGAAACAGAGAACCCCTCGCCACG 169
DB 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGTGCCCTCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATCGAGTCT 229
DB 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGCTGGACCAAGCTCCAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGGC 289
DB 61 ArgCysThrLysLeuGluIyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAACGTTCCCTCCAGGGGACGACACAGCTGGCCGCCAACCCGAGACTAC 349
DB 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCGTCAACCCGAGAGAGAGAGGCGCGATGGGACCCAGCTGGACCG 409
DB 101 AspAspArgArgGlnProArgGluGluGlyArgTrpGlyProAlaGlyPro 120
QY 410 AGGAGCGGTCAAAGAGAAGAGACTGGAGACACACCAAGAGAAGATTGGAGGCCACCAAGT 469
```

```
DB 121 ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
QY 470 CATCAGCAGCCACGAAATAGCCCGAAGGAAGAGAGAGAAACAAGAGTGGGAAACA 529
DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
QY 530 CCAGTAGGCATGTCAGGGAGAAACATCTCGGAACAACCCCTTCTACTTCCCGTCAAGG 589
DB 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCCGCTACGGGAACCAAAACCGTAGGATCCGGGTCCTCGCAGAGTTTGAC 649
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgGlyArgValGluGlnArgPheAsp 200
QY 650 CAAAGGTCAAGCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCCAGGCCAAA 709
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTTGTCTTCCCAAGCACGCTGATGCTGATAACATCCTCTTATCCAGCAA 769
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTCACCGTAGCAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAG 829
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCATGCACTCAGAAATCCCATCCCGTTTCATTTCTCATCTTGAACCCCATGACAAC 889
DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCCTCAGAGTAGCTAAATCTCCATGCCCGTTAAACACACCCGCCAGTTTGAAGAT 949
DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCGCAGCAGCCGACCAATCATCTACTTGCAGGGCTTCAGAGGAATACG 1009
DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGluPheSerArgAsnThr 320
QY 1010 TTGAGAGCGCCTTCAATCCGGAATTAATGATGATACGAGGGTGTCTTAGAAGAGAT 1069
DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGGTGACCAAGAGAGAGAGCGGAGCGGATGAGTACTCGGAGTAGTAGAGAAC 1129
DB 341 AlaGlyGlyGluGlnGluArgGlyGlnArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGAGTAGTAGTCAAAGTGTCAAAGGACACGTTGAAAGAACTTACTAAGCAGCT 1189
DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCGCTCAAAGAAAGCTCCGGAAGAGAGGGAGATATCACCAACCCCACTCACTTG 1249
DB 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGAGCGGAGCCCGCATCTTTCTAAACAACCTTGGGAAGTATTTCAGGTGAAGCCAGAC 1309
DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAACCCCGAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCTCGTCCGTCACAAA 1429
DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAlaAsnLys 460
QY 1430 GGAACCTGGAAAACCTTCAACTCGTGCTGTAAAGAAAAAGAGCAACAACAGAGGGGACCGCGG 1489
DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnArgGlyArgArg 480
QY 1490 GAAGAAAGAGGAGCAGAACGAAAGAGAGGAGGAAAGTAAACAGAGAGGTGCGTAGGTAC 1549
DB 481 GluGluGluGluAspGluGluGluGluGluGluGluGluGluGluGluValArgGlyTyr 500
```

QY 1550 ACAGCGAGGTTGAAGCAAGCGCATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
 Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
 QY 1610 AAGCTTCTCCGAATCCATCTGCTGGTTCGGTATCAACGCTGCAAAACACACACAGA 1669
 Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg 540
 QY 1670 ATCTCTCTGCGAGTCAAGACAAATGTATAGACCATAGAGAGCAAGCAAGGAT 1729
 Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
 QY 1730 TTAGCATTCCTCGGTCGGTGCAACAAAGTTGGAAGCTCATCAAAACACAGAGGAATCT 1789
 Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
 QY 1790 CACTTTGTGAGTCTGCTCTCAATCTCAATCTCAATCTCGTCTCGTCTGAGAAAGAG 1849
 Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
 QY 1850 TCTCCTGAGAAAGAGGATCAAGAGGAGGAAACCAAGGAGGAGGATCCACTCTCTTCA 1909
 Db 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
 QY 1910 ATTTTGAAGCTTTTAAAC 1927
 Db 621 IleLeuLysAlaPheAsn 626

RESULT 2

US-09-106-872A-17
 ; Sequence 17, Application US/09106872A
 ; Patent No. 6486311
 ; GENERAL INFORMATION:
 ; APPLICANT: Burks Jr., A. Wesley
 ; APPLICANT: Stanley, J. Steven
 ; APPLICANT: Cockrell, Gael
 ; APPLICANT: King, Nina E.
 ; APPLICANT: Sampson, Hugh A.
 ; APPLICANT: Helm, Ricki M.
 ; APPLICANT: Bannon, Gary A.
 ; TITLE OF INVENTION: Peanut Allergens and Methods
 ; FILE REFERENCE: HS 103 CIP
 ; CURRENT APPLICATION NUMBER: US/09/106,872A
 ; CURRENT FILING DATE: 1999-06-29
 ; PRIOR APPLICATION NUMBER: PCT/US96/15222
 ; PRIOR FILING DATE: 1996-09-23
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 17
 ; LENGTH: 335
 ; TYPE: PRT
 ; ORGANISM: Arachis hypogaea
 US-09-106-872A-17

Alignment Scores:
 Pred. No.: 2,28e-157 Length: 335
 Score: 1700.00 Matches: 335
 Percent Similarity: 99.4% Conservative: 0
 Best Local Similarity: 99.4% Mismatches: 0
 Query Match: 47.4% Indels: 2
 DB: 2 Gaps: 1

US-10-728-323-1 (1-2032) x US-09-106-872A-17 (1-335)

QY 917 ATGCCCGTTAAACACACCCGGCCAGTTTGGAGTATCTCCCGCGAGGAGCGGACACAA 976
 Db 1 MetProValAsnThrProGlyGlnPheGluAspPhePheProAlaSerSerArgAspGln 20
 QY 977 TCATCTACTTCAGGCGTTTCAGCAGGAATACGTTGGAGCGCGCTTCAATCGGAATTC 1036
 Db 21 SerSerTyLeuGlnGlyPheSerArgAsnThrLeuGluAlaAlaPheAsnAlaGluPhe 40

QY 1037 AATGAGATACGAGGCTGCTGTTAGAGAGAAATGCAGAGGTGACAAAGAGAGAGAGG 1096
 Db 41 AsnGluIleArgArgValLeuLeuGluAsnAlaGlyGlyGluGlnGluArgGly 60
 QY 1097 CAGAGCGCATGAGTACTCGAGTACTGAGAACAAATGAAGGAGTATAGTCAAAGTGTCA 1156
 Db 61 GlnArgArgTrpSerThrArgSerSerGluAsnAsnGluGlyValIleValLysValSer 80
 QY 1157 AAGGAGCACGTTTGAAGAACTTACTAAGCACGCTAAATCCGCTCTCAAGAAAGGCTCCGAA 1216
 Db 81 LysGluHisValGluGluLeuThrLysHisAlaLysSerValSerLysLysGlySerGlu 100
 QY 1217 GAAGAGGAGATATACCAACCAATCAACTTTGAGAGAAGCGAGCCGATCTTTCTAAC 1276
 Db 101 GluGluGlyAspIleThrAsnProIleAsnLeuArgGluGlyGluProAspLeuSerAsn 120
 QY 1277 AACTTTGGGAGTTATTTGAGGTGAAGCCAGACAGAGAGAGAGAGAGAGAGAGAGAG 1336
 Db 121 AsnPheGlyLysLeuPheGluValLysProAspLysLysAsnProGlnLeuGlnAspLeu 140
 QY 1337 GACATGATGCTCACCTGTGTAGAGATCAAGAAAGAGGCTTTTGTGCTCCACACTTCAAC 1396
 Db 141 AspMetMetLeuThrCysValGluIleLysGluGlyAlaLeuMetLeuProHisPheAsn 160
 QY 1397 TCAAGGCGCATGTTATCGTCTCTCAACAAAGAACTGGAACCTTGAACCTCGTGGCT 1456
 Db 161 SerLysAlaMetValIleValValAsnLysGlyThrGlyAsnLeuGluLeuValAla 180
 QY 1457 GTAAGAAAGAGCAACACAGAGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 1516
 Db 181 ValArgLysGluGlnGlnArgGlyArgArgGluGluGluGluAspGluAspGluGlu 200
 QY 1517 GAGGAGGGAAGTAACACAGAGAGGCTAGGTACACAGGAGGTTTGAAGAGGCGCATGTG 1576
 Db 201 GluGluGlySerAsnArgGluValArgArgTyrThrAlaArgLeuLysGluGlyAspVal 220
 QY 1577 TTCATATGCCAGCAGCTCATCCAGTACCATCAACGCTTCTCCGAACTCCATCTGCTT 1636
 Db 221 PheIleMetProAlaAlaHisProValAlaIleAsnAlaSerSerGluLeuHisLeuLeu 240
 QY 1637 GCTTTGCGTATCAACGCTGMAAACCAACACAGAACTTCTTGCAGGTGATAGGACAT 1696
 Db 241 GlyPheGlyIleAsnAlaGluAsnAsnHisArgIlePheLeuAlaGlyAspLysAspAsn 260
 QY 1697 GTGATAGACCATAGAGAAGCAAGCAAGGATTTTAGCATTTCCCTGGGTGCGGTGAACAA 1756
 Db 261 VallieAspGlnIleGluLysGlnAlaLysAspLeuAlaPheProGlySerGlyGluGln 280
 QY 1757 GTTGAGAGCTCATCAAAACACAGAAAGAAATCTCACTTTGTGAGTGTCTGCTCTCATCT 1816
 Db 281 ValGluLysLeuIleLysAsnGlnLysGluSerHisPheValSerAla-----GlnSer 298
 QY 1817 CAATCTCAATCTCGTCTCTCCAGAAAGAGTCTCTGAGAAAGAGTCTCTGAGAAAGAGATCAAGAGGAG 1876
 Db 299 GlnSerGlnSerProSerSerProGluLysGluSerProGluLysGluAspGlnGluGlu 318
 QY 1877 GAAACCAAGGAGGAGGCTCCACTCTCTTCAATTTTGAAGGCTTTTAAAC 1927
 Db 319 GluAsnGlnGlyGlyLysGlyProLeuLeuSerIleLeuLysAlaPheAsn 335

RESULT 3

US-07-955-905A-24
 ; Sequence 24, Application US/07955905A
 ; Patent No. 5770433
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: RECOMBINANT 47 AND 31 KD COCOA PROTEINS AND
 ; TITLE OF INVENTION: PRECURSOR
 ; NUMBER OF SEQUENCES: 28
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 605 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..605
; OTHER INFORMATION: /note= "Vicillin from G. max"
;
US-07-955-905A-24

Alignment Scores:
Pred. No.: 5,47e-116 Length: 605
Score: 1279.00 Matches: 285
Percent Similarity: 59.2% Conservative: 107
Best Local Similarity: 43.1% Mismatches: 174
Query Match: 35.7% Indels: 96
DB: Gaps: 18

US-10-728-323-1 (1-2032) x US-07-955-905A-24 (1-605)
Qy 50 ATCAGAGGAGGAGTTTCCACTGATGCTGTGTAGGGATCCTGTGCTGGCTTCAGTT 109
Db 2 MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyLeuValPheLeuAlaSerVal 19
Qy 110 TCTCAACACGATGCCAAGTCATCACTTACCAGAGAGAAAACAGAGAACCCCTCGCCACG 169
Db 20 SerValSerPheGlyLeaLa-----TyrTrpGluLys---GluAsnProLysHisAsn 36
Qy 170 AGTGTCTCCAGATTGTTCACAGAAACCGGATGACTTTGAAGCAAAAGGATCGAGTCT 229
Db 37 LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 56
Qy 230 CGCTGCACCAACGTCGATGATGAT-----CCTCGTTGT 262
Db 57 ArgCysAsnLeuLeuLysValGluLysGluGluCysGluGluGlyGluIleProArg 75
Qy 263 GTCTATGATCTCGAGGACACACTGGCACCAACCAACCAAGTTCCTCCAGGGGAGCGG 322
Db 76 -----ProArgProArg-----ProGlnHisProGlu 84
Qy 323 ACACGTGGCGGCAACCCGGAGCTACGATGATGAC----- 358
Db 85 ArgGluProGlnGlnProGlyGluLysGluAlaAspGluAspGluGlnProArgProIle 104
Qy 359 -----CGCGCTCAACCCCGAAGAGAG-----GAAGGAGGCCGA 391
Db 105 ProPheProArgProGlnProArgGlnGlnGluHisGluGlnArgGluGluGlnGlu 124
Qy 392 TGG-----CGACCACTGACCGAGGAGCGCTGAAAGAGAGAA 430
Db 125 TrpProArgLysGluLysArgGlyGluLysGlySerGluGluAspGluAspGlu 144
Qy 431 GACTGAGACAAACAGAGAAGAT-----TGGAGGCGACCAAGTCTATCAGCAGCCACGG 484
Db 145 AspGluGluGlnAspGluArgGlnPheProPheProArgProProHisGlnLysGluGlu 164
Qy 485 AAAATAAGGCCCGAAGAGAGAGAGAAAGAGTGGGGAAACACCAAGGTAGCCATGTG 544
Db 165 ArgAsnGluGluAspGluAspGluGlnGlnArgGluSerGluSerGluAsp 184
Qy 545 AGGGAA-----GAAACATCTCGGAACAAACCTTTCTACTTCCCGTCAAGCGGTTTACG 598
Db 185 SerGluLeuArgArgHisLysAsnLysAsnProPheLeuPhePheSerAsnArgPheGlu 204

```

Db 540 AlaGlySerGlnAspAsnValIleSerGlnIleProSerGlnValGlnGluLeuAlaPhe 559
| | | | | : : : | | | | | | | | | | | | | | | : : : : : | | | | |
Qy 1739 CCTGGCTCGGTGAACAACTTGAGAGCTCATCAAAACCCAGAACGGAATCTCACTTGTG 1798
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 560 ProGlySerAlaGlnAlaValGluLysLeuLeuLysAsnGlnArgGlySerTyrPheVal 579
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 1799 AGTGCTCGCTCAATCTCAATCTCAATCTCCGCTCCTCCTGAGAAAGAGTCTCCTGAG 1858
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 580 AspAlaGlnProLysLys 585
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 1859 AAGAGGATCAAGAGAGGAAACCAAGAGGAGGAGGTCCTCTTCAATTTTGAAG 1918
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 586 -----LysGluGluGlyAsnLysGlyArgLysGlyProLeuSerSerIleLeuArg 602
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 1919 GCTTTT 1924
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 603 AlaPhe 604
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |

RESULT 4

US-07-955-905A-25
; Sequence 25 Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND
; TITLE OF INVENTION: PRECURSOR
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 571 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Pseum sativum
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..571
; OTHER INFORMATION: /note= "Conviciilin from P. sativum"
US-07-955-905A-25

Alignment Scores:

Pred. No.: 7.88e-115 Length: 571
Score: 1267.00 Matches: 276
Percent Similarity: 56.7% Conservative: 98
Best Local Similarity: 41.8% Mismatches: 146
Query Match: 35.3% Indels: 140
DB: 1 Gaps: 13

US-10-728-323-1 (1-2032) x US-07-955-905A-25 (1-571)

Qy 68 CCAGTGTGCTGTAGGGATCCTTGTCTGGCTTCAGTTTCTGCAACGATGCCAAG 127
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 10 ProLeuLeuPheLeuGlyIleIlePheLeuAlaSerValCysValThrTyrAlaAsn 29
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 128 TCATCACTTACAGAAGAAACAGAGAACCCCTGCGCCAGAGGTGCTCCAGAGTTGT 187
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 29 ----- 29
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 188 CAACAGAACCGGATGACTTTGAAGCAAAAGGCATGCGAGTCTCGCTGCACCAAGCTCGAG 247
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 29 ----- 29
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |

Qy 248 TATGATCCTCGTGTGTCTATGATCCTCGAGGACACACTGGCACCACCACCAACGCTTCC 307
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 30 -----TyrAspGluGlySerGluThrArgVal----- 38
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 308 CCTCAGGGGACCGGACACAGTGGCGCCCAACCCGAGAGACTACGATGATGATCAGCCCGTCAA 367
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 39 ---ProGlyGlnArgGluArgGlnGluGlyGluLysGluLysArgHis--- 56
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 368 CCCCAAGAGAGAGGAGGCGGATGGGACCCAGCTGGACCCGAGGAGCGT----- 418
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 57 -----GlyGluTrpArgProSerTyrGluLysGluGluHisGluGlu 70
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 418 ----- 418
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 71 GluLysGlnLysTyrArgTyrGlnArgGluLysLysGluGlnLysGluValGlnProGly 90
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 419 -----GAAAGAGAGAACTGGAGACACCAACCAAGAGAGATTGGAGG----- 460
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 91 ArgGluArgTrpGluArgGluAspGluGlnValGlnGluGluTrpArgGlySer 110
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 461 -----CGACCAAGTCATCAGCAGGCCACCG 484
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 111 GlnArgArgGluAspProGluGluArgAlaArgLeuArgHisArgGluGluArgThrLys 130
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 485 AAAATAAGGCCCGAAGAGAGAGAGAGAGAGAGTGGGGAACACCAAGGTAGCCATGTG 544
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 131 ArgAspArgHisGlnArgGluGlyGluGluGlu-----GluArgSer 145
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 545 AGGGAAGAAACATCTCGGAACCAACCTTTCTACTTCCCGTCAAGCGGTTTAGCACCCTG 604
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 146 SerGluSerGlnGluHisArgAsnProPheLeuPheLysSerAsnLysPheLeuThrLeu 165
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 605 TACGGGAACCAACCGTAGGATCCGGGTCCTCGAGAGGTTTGACCAAGGTCAAGGCCAG 664
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 166 PheGluAsnGluAsnGlyHisIleArgArgLeuGlnArgPheAspLysArgSerAspLeu 185
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 665 TTTTCAAGATCTCCAGAAATCACCGTATTGTGAGATCGAGGCCAAACCTAACACTCTTGT 724
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 186 PheGluAsnLeuGlnAsnTyrArgLeuValGluTyrArgAlaLysProHisThrIlePhe 205
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 725 CTTCCCAAGCAGCTGATGCTGATTAACATCCTCTGTTATCCACCAAGGCCAAGCCCGTG 784
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 206 LeuProGlnHisIleAspAlaAspLeuIleLeuValLeuAsnGlyLysAlaIleLeu 225
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 785 ACCGTAGCAAAATGCAATAACAGAAAGAGCTTTAATCTTGACGAGGCCATGCACCTCAGA 844
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 226 ThrValLeuSerProAsnAspArgAsnSerTyrAsnLeuGluArgGlyAspThrLys 245
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 845 ATCCCATCCGTTTCATTTCTTACATCTTGAAACCCGCATGACACCAAGAACCTCAGAGTA 904
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 246 IleProAlaGlyThrThrSerTyrLeuValAsnGlnAspAspGluGluAspLeuArgVal 265
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 905 GCTAAATCTCCATCCCGTTTAAACACACCCCGCCAGTTTGAGGATTTCTTCCCGGCGAGC 964
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 266 ValAspPheValIleProValAsnArgProGlyLysPheGluAlaPhe-----GlyLeu 283
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 965 AGCCGAGACCAATCATCTTCTTACATCTTGAAACCCGCATGACACCAAGAACCTCAGAGTA 1024
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 284 SerGluAsnLysAsnGlnTyrLeuArgGlyPheSerLysAsnIleLeuGluAlaSerLeu 303
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 1025 AATCGGGAATCAATGAGATACGAGGAGTGTGTTAGAGAGAAATGCAGAGGTGAGCAA 1084
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 304 AsnThrLysTyrGluThrIleGluLysValLeuLeuGluGlnGluLysLysProGln 323
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 1085 GAGGAGAGAGGCGACAGGCGATGAGTACTCGAGTAGTAGGAAACAAATGAGAGGTGATA 1144
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 324 GlnLeuArgAspArgLysArg-----ThrGlnGlnGlyGluGluArgAsp---AlaIle 340
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 1145 GTCAAGTGTCAAAGGACGACGCTTGAAGAACTTACTTAAGCAGCTAAATCCGCTCTCAAAG 1204
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 341 IleLysValSerArgGluGlnIleGluGluLeuArgLysLeuAlaLysSerSerLys 360
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 1205 AAAGGCTCCGAAGAGAGGAGATATACCAACCCCAATCACTTGTGAGAGAGCGAGCCC 1264
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |

```
Db      361  LysSerLeuProSerGluPheGlu-----ProPheAsnLeuArgSerHisLysPro 377
      1265  GATCTTCTTAACAACATTTGGGAAGTTATTGAGTGAAGCCAGACAGAAAGAACCCCCAG 1324
      378  GluTyrSerAsnLysPheGlyLysLeuPheGluLeuThrProGluLysLysTyrProGln 397
      1325  CTTTCAGGACTGACATGATGCTCACCCTGTGTAGAGATCAAGAAGAGCTTTGTGCTC 1384
      398  LeuGlnAspLeuAspIleLeuValSerCysValGluIleAsnLysGlyAlaLeuMetLeu 417
      1385  CCACACTTCAACTCAAGGCCATGGTTATCGTCGTCTCAACAAAGAACTGGAACCTT 1444
      418  ProHisTyrAsnSerArgAlaIleValValLeuLeuValAsnGluGlyLysGlyAsnLeu 437
      1445  GAATCTCGTCTGTAAAGAAAGAGCAACACAGAGGGGACGGCGGAAGAAGAGGAGGAC 1504
      438  GluLeuLeuGlyLeuLysAsnGluGlnGluArg----- 449
      1505  GAAGACGAAGAAGAGAGGAAGTAACAGAGAGGTGCGTAGGTACACAGGAGGTGAAG 1564
      450  GluAspArgLysGluArg-----AsnAsnGluValGlnArgTyrGluAlaArgLeuSer 467
      1565  GAAGCGCATGTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCCCTCCGAA 1624
      468  ProGlyAspValValIleIleProAlaGlyHisProValAlaIleSerAlaSerAsn 487
      1625  CTCATCTGTGCTGGCTTCGGTATCAACGCTGAAACAAACACACAGAAATCTTCTTCGAGGT 1684
      488  LeuAsnLeuLeuGlyPheGlyIleAsnAlaLysAsnAsnGlnArgAsnPheLeuSerGly 507
      1685  GATAAGACAATGTATAGACCAAGATAGAGAAGCAAGCAAGGATTTAGCATTCCTCGGG 1744
      508  SerAspAspAsnValIleSerGlnIleGluAsnProValLysGluLeuThrPheProGly 527
      1745  TCGGGTCAACAAGTGTGAGAGCTCATCAAAACCAAGAAATCTCACTTTGTGAGTGT 1804
      528  SerSerGlnGluValAsnArgLeuIleLysAsnGlnLysGlnSerHisPheAlaSerAla 547
      1805  CGTCTCAATCTCAATCTCAATCTCCGTCGTCTCCTCGAGAAAGAGTCTCCTGAGAAAGAG 1864
      548  GluPro----- 550
      1865  GATCAAGAGAGGAAACCAAGAGGAGGAGGAGGTCCTCTTCAATTTTGAAGGCTTTT 1924
      551  GlnLysGluGluGluSerGlnArgLysArgSerProLeuSerSerValLeuAspSerPhe 570

RESULT 5
US-10-245-227D-91
; Sequence 91, Application US/10245227D
; Patent No. 6936696
; GENERAL INFORMATION:
; APPLICANT: Monsanto Company
; APPLICANT: Rapp, William D.
; APPLICANT: Peng, JieXin
; APPLICANT: Nadiq, Gautham
; APPLICANT: Venkatesh, Tyamagondlu
; TITLE OF INVENTION: Enhanced Protein and Methods of Their Use
; FILE REFERENCE: REN-00-087 US
; CURRENT APPLICATION NUMBER: US/10/245,227D
; CURRENT FILING DATE: 2002-09-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 91
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Glycine max
US-10-245-227D-91

Alignment Scores:
Pred. No.: 1.69e-105 Length: 449
Score: 1171.00 Matches: 239
Percent Similarity: 70.4% Conservative: 87
```

```
Best Local Similarity: 51.6% Mismatches: 83
Query Match: 32.7% Indels: 54
DB: 2 Gaps: 10

US-10-728-323-1 (1-2032) x US-10-245-227D-91 (1-449)

QY      542  GTGAGGAGAAACATCTCGGAACACCCCTTTCTACTTC--CCGTCAAGGGGTTTAGC 598
Db      36  ValArgGluAsp-----GluAsnAsnProPheTyrLeuArgSerSerAsnSerPheGln 53
QY      599  ACCCGCTACGGGAACCAAAACGGTAGGATCCGGGCTCTCGAGAGGTTTACCAAAAGTCA 658
Db      54  ThrLeuPheGluAsnGlnAsnGlyArgIleArgLeuIleGlnArgPheAsnLysArgSer 73
QY      659  AGCGAGTTTCAGAAATCCAGAAATCACCGTATTTGTCAGATCGAGGCCAACCTAACACT 718
Db      74  ProGlnLeuGluAsnLeuArgAspTyrArgIleValGlnPheGlnSerLysProAsnThr 93
QY      719  CTTGTCTTCCCAAGCACCGCTGATGCTGATAACATCCTTTGTTATCCAGCAAGGGCAAGCC 778
Db      94  IleIleLeuProHisHisAlaAspAlaAspPheIleIleIleIleLeuSerGlyArgAla 113
QY      779  ACCGTACCGTAGCAAAATGGCAATACAGAAAAGAGCTTTAATCTTCACGAGGGCCATGCA 838
Db      114  IleLeuThrLeuValAsnAsnAspAspArgAspSerTyrAsnIleHisProGlyAspAla 133
QY      839  CTCAGAAATCCCATCCCGTTTCATTTCTTACATCTTGAACCGCCATCAACACCAAGACCTC 898
Db      134  GlnArgIleProAlaGlyThrThrTyrTyrIleIleAsnProHisAspHisGlnAsnLeu 153
QY      899  AGAGTAGCTAAATCTCCATGCCGTTTAAACACACACCGCCAGTTTCAGGATTTCTTCCCG 958
Db      154  LysIleIleLysIleAlaIleProValAsnLysProGlyArgTyrAspAspPhePheLeu 173
QY      959  GCGAGCAGCGCAGACCAATCATCTACTTTCGAGGGCTTCAGCAGGAATACGTTGGAGGCC 1018
Db      174  SerSerThrGlnAlaGlnGlnSerTyrLeuGlnGlyPheSerHisAsnIleLeuGluThr 193
QY      1019  GCCTTCAATCGGAATTCATAGATACGAGAGGTCCTGTTAGACAGAGATCGCAGAGGT 1078
Db      194  SerPheHisSerGluPheGluGluIleAsnArgValLeuLeu-----Gly 208
QY      1079  GAGCAGAGAGAGAGGGCAGAGCGATGGAGTACTCGGAGTAGTCAGAACAAATGAAGGA 1138
Db      209  GluGluGluGluGlnArgGlnGln----- 218
QY      1139  GTGATAGTCAAAAGTGTCAAAGGACGTTTGAAGAACTTACTAAGCAGCGTAAATCCGTC 1198
Db      219  ValIleValGluLeuSerLysGluGlnIleArgGlnLeuSerArgArgAlaLysSerSer 238
QY      1199  TCAAAGAAA-----GGCTCCGAGAGAGGGGAGATATCACCAACCAATCACTTTCAGA 1252
Db      239  SerArgLysThrIleSerSerGluAspGlu-----ProPheAsnLeuArg 253
QY      1253  GAAGGCGAGCCGATCTTCTTAACTTTGGAAAGTTATTTGAGGTGAAGCCAGCAAG 1312
Db      254  SerArgAsnProIleTyrSerAsnAsnPheGlyLysPheGluIleThrProGlu--- 272
QY      1313  AAGAAACCCCGAGCTTCAGGACCTGGACATGATCCTCACCTGTGTAGAGATCAAGAGGA 1372
Db      273  LysAsnProGlnLeuArgAspLeuAspIlePheLeuSerSerValAspIleAsnGluGly 292
QY      1373  GCTTTGATGCTCCACACTTCACTCAAGGCCATGGTTATCTGCTCGTCAACAAGGA 1432
Db      293  AlaLeuLeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGly 312
QY      1433  ACTGGAAACCTTGAACTCGTGGCTGTAAAGAAAAGAGCAACACAGAGGGGACCGCGGAA 1492
Db      313  AspAlaAsnIleGluLeuValGlyIle---LysGluGlnGlnGlnLysGlnLysGlnGlu 331
QY      1493  GAAGAGAGGACGAAGACGAAGAGAGGAGGAGGTAACTAACAGAGAGGTGCTAGTACACA 1552
Db      332  GluGluPro-----LeuGluValGlnArgTyrArg 341
```


; APPLICANT: Peng, Jixen
; APPLICANT: Nadig, Gautham
; APPLICANT: Venkatesh, Tyamagondlu
; TITLE OF INVENTION: Enhanced Protein and Methods of Their Use
; FILE REFERENCE: REN-00-087 US
; CURRENT APPLICATION NUMBER: US/10/245,227D
; CURRENT FILING DATE: 2002-09-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Glycine max
US-10-245-227D-1

Alignment Scores:
Pred. No.: 1,01e-104 Length: 439
Score: 1163.00 Matches: 238
Percent Similarity: 70.2% Conservative: 87
Best Local Similarity: 51.4% Mismatches: 84
Query Match: 32.4% Indels: 54
DB: 2 Gaps: 10

US-10-728-323-1 (1-2032) x US-10-245-227D-1 (1-439)

```
QY 542 GTGAGGGAAGAAACATCTCGGAACACCCCTTTCTACTTC---CCGTCAAGCGGTTTAGC 598
DB 26 ValArgGluAsp-----GluAsnAsnProPheTyrLeuArgSerSerAsnSerPheGln 43
QY 599 ACCCGTACGGGAACCAAAACGGTAGGATCGGGTCTCGAGAGGTTTGACCAAGGTCA 658
DB 44 ThrLeuPheGluAsnGlnAsnGlyArgIleArgLeuGlnArgPheAsnLysArgSer 63
QY 659 AGGCGATTTCAGAACTTCAGAAATCACCGTATTGTGCAGATCGAGGCCAACTACACT 718
DB 64 ProGlnLeuGluAsnLeuArgAspTyrArgIleValGlnPheGlnSerLysProAsnThr 83
QY 719 CTTGTTCTTCCCAAGCACGCTGATGCTGATAACATCTCTTGTATTCCAGCAAGCGCAAGCC 778
DB 84 IleLeuLeuProHisHisAlaAspAlaAspPheLeuLeuPheValLeuSerGlyArgAla 103
QY 779 ACCGTACCGTACCAATGCGCAATACAGAAAGAGCTTTAATCTTGACGAGGCCATGCA 838
DB 104 IleLeuThrLeuValAsnAsnAspAspArgAspSerTyrAsnLeuHisProGlyAspAla 123
QY 839 CTCAGAAATCCATCCCGTTTCATTCTCATCTTCAACCGCATGACACACCAACGACCTC 898
DB 124 GlnArgIleProAlaGlyThrThrTyrTyrLeuValAsnProHisAspHisGlnAsnLeu 143
QY 899 AGAGTAGTAAATCTCCATGCCCGTTAAACACACCGCGCAGTTTGAGGATTTCTTCCCG 958
DB 144 LysIleLeuLeuAlaIleProValAsnLysProGlyArgTyrAspAspPhePheLeu 163
QY 959 GCGAGCAGCGGACCAATCATCTCTTGCAGGGCTTCAGAGGAATACGTTGGAGGCC 1018
DB 164 SerSerThrGlnAlaGlnGlnSerTyrLeuGlnGlyPheSerHisAsnIleLeuGluThr 183
QY 1019 GCCTCAATCGGAATTCATAGATACGAGGGTCTCTAGAAGAGATCGCAGAGGT 1078
DB 184 SerPheHisSerGluPheGluGluIleAsnArgValLeuLeu-----Gly 198
QY 1079 GAGCAAGAGAGAGCGGCGATGGAGTACTCGGAGTAGTGAGAACAAATGAAGA 1138
DB 199 GluGluGluGluGlnArgGlnGln-----GluGly 208
QY 1139 GTGATAGTCAAGTCAAGAGGACGCTTGTAAGAACTTACTAAGCACCGCTAAATCCGTC 1198
DB 209 ValIleValGluSerLysGluGlnIleArgGlnLeuSerArgArgAlaLysSerSer 228
QY 1199 TCAAGAAA-----GGCTCCGAGAGAGGGAGATATCACCAACCAATCACTTGAGA 1252
DB 229 SerArgLysThrIleSerSerGluAspGlu-----ProPheAsnLeuArg 243
```

RESULT 9

US-10-245-227D-89
; Sequence 89, Application US/10245227D
; Patent No. 6936696
; GENERAL INFORMATION:
; APPLICANT: Monsanto Company
; APPLICANT: Rapp, William D.
; APPLICANT: Peng, Jixen
; APPLICANT: Nadig, Gautham
; APPLICANT: Venkatesh, Tyamagondlu
; TITLE OF INVENTION: Enhanced Protein and Methods of Their Use
; FILE REFERENCE: REN-00-087 US
; CURRENT APPLICATION NUMBER: US/10/245,227D
; CURRENT FILING DATE: 2002-09-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 89
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Glycine max
US-10-245-227D-89

Alignment Scores:

Pred. No.: 1.03e-104 Length: 449
Score: 1163.00 Matches: 238

```
Percent Similarity: 70.2%      Conservative: 87
Best Local Similarity: 51.4%    Mismatches: 84
Query Match: 32.4%             Indels: 54
DB: 2                           Gaps: 10

US-10-728-323-1 (1-2032) x US-10-245-227D-89 (1-449)

QY 542 GTGAGGAGAAACATCTCGGAACACCCCTTTCTACTTC---CCGTCAAGGCGGTTTAGC 598
Db 36 ValArgGluAsp-----GluAsnAsnProPheTyrLeuArgSerSerAsnSerPheGln 53
QY 599 ACCCGGTACGGGAACCAAAACGGTAGATCCGGTCTCTCAGAGGTTTGACCAAAAGGTCA 658
Db 54 ThrLeuPheGluAsnGlnAsnGlyArgIleArgLeuGlnArgPheAsnLysArgSer 73
QY 659 AGGCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCAGGCCCAACCTAACACT 718
Db 74 ProGlnLeuGluAsnLeuArgAspTyrArgIleValGlnPheGlnSerLysProAsnThr 93
QY 719 CTTGTTCTCCCAAGCACGCTGATGCTGATAACATCCTTGTATTCCAGCAAGGGCAAGCC 778
Db 94 IleLeuLeuProHisHisAlaAspPheLeuLeuPheValLeuSerGlyArgAla 113
QY 779 ACCGTGACCGTAGCAATGCGAATAACAGAAAGAGCTTTAAATCTTGACGAGGGCCATGCA 838
Db 114 IleLeuThrLeuValAsnAsnAspAspArgAspSerTyrAsnLeuHisProGlyAspAla 133
QY 839 CTCAGAAATCCCATCCGGTTTCATTCTCATCTTGAACCGCCATGACAAACCAAGAACCTC 898
Db 134 GlnArgIleProAlaGlyThrThrTyrLeuValAsnProHisAspHisGlnAsnLeu 153
QY 899 AGAGTAGCTAAATCTCCATGCCGTTTAAACACACCCGGCCAGTTTGAGGATTTCTCCCG 958
Db 154 LysIleIleLysLeuAlaIleProValAsnLysProGlyArgTyrAspAspPhePheLeu 173
QY 959 GCGAGGACGCGACCAATCATCTCTTCGCGGCTTCAGCAGGAATACGTTGGAGGCC 1018
Db 174 SerSerThrGlnAlaGlnSerTyrLeuGlnGlyPheSerHisAsnIleLeuGluThr 193
QY 1019 GCCTTCAATCGGAATTCATAGATACGAGGAGTCTGTTAGAGAGATGACGAGAGGT 1078
Db 194 SerPheHisSerGluPheGluGluIleAsnArgValLeuLeu-----Gly 208
QY 1079 GACCAAGAGAGAGAGGGGACAGCGATGAGTACTCGGAGTAGTGAGAACATGAGGA 1138
Db 209 GluGluGluGluGlnArgGlnGln-----GluGly 218
QY 1139 GTGATAGTCAAAGTGTCAAAGGACGTTGAAGAACTTACTAAGCACGCTAAATCCGTC 1198
Db 219 ValIleValGluLeuSerLysGluGlnIleArgGlnLeuSerArgArgAlaLysSerSer 238
QY 1199 TCAAAGAAA-----GCCTCCGAAGAGAGGGAGATATACCAACCAATCAACTTGAGA 1252
Db 239 SerArgLysThrIleSerSerGluAspGlu-----ProPheAsnLeuArg 253
QY 1253 GAAGGCGAGCCCATCTTTCTCAACATTTGGGAAGTTATTTCAGGTGAGCCAGACAAG 1312
Db 254 SerArgAsnProIleTyrSerAsnAsnPheGlyLysPheGluIleThrProGlu--- 272
QY 1313 AAGAACCCCGAGCTTCAGGACCTGGACATGATGCTACCTGTGTAGAGATCAAAGAGGA 1372
Db 273 LysAsnProGlnLeuArgAspLeuAspIlePheLeuSerSerValAspIleAsnGluGly 292
QY 1373 GCTTTGATGTCCTCCACATTTCAACTCAAAGCCATGTTATCGTCGTCGTCACAAAGGA 1432
Db 293 AlaLeuLeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGly 312
QY 1433 ACTGGAACCTTGAACCTCGGTGTGAAGAAAAGAGCAACAAAGAGGGGACCGCGGAA 1492
Db 313 AspAlaAsnIleGluLeuValGlyIle---LysGluGlnGlnLysGlnLysGlnGlu 331
QY 1493 GAAGAGGAGGACGAAGACGAAGAAGAGGAGGGAAGTAACAGAGAGGTGCGTAGGTACACA 1552
Db 1552
```

```

Db 332 GluGluPro-----LeuGluValGlnArgTyrArg 341
QY 1553 GCGAGGTTGAAGAAAGCGCATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAAC 1612
Db 342 AlaGluLeuSerGluAspAspValPheValIleProAlaAlaTyrProPheValValAsn 361
QY 1613 GCTTCTCCGAACCTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACACCAACCAACAATC 1672
Db 362 AlaThrSerAsnLeuAsnPheLeuAlaPheGlyIleAsnAlaGluAsnGlnArgAsn 391
QY 1673 TTCTTTCAGAGTGATAAGACAAATGTGATAGACCCAGATAGAGAAGCAAGCAAGGATTTA 1732
Db 382 PheLeuAlaGlyGluLysAspAsnValValArgGlnIleGluArgGlnValGlnGluLeu 401
QY 1733 GATTCCTCCGTCGGTGGTGAAACAAAGTTGAGAAAGCTCATCAAAACACAGAAAGGATCTC 1792
Db 402 AlaPheProGlySerAlaGlnAspValGluArgLeuLeuLysGlnArgGluSerTyr 421
QY 1793 TTTGTGAGTCTGCTGCTCCTCAATCTCAATCTCAATCTCCTCGTCGTCCTCGAGAAAGAGTCT 1852
Db 422 PheValAspAlaGlnProGln----- 428
QY 1853 CCTGAGAAAGAGGATCAAGAGGAGGAAACCAAGGAGGAGGAGGTCCTCTTCAATT 1912
Db 429 -----GlnLysGluGluGlySerLysGlyArgLysGlyProPheProSerIle 444
QY 1913 TTGAAGGCT 1921
Db 445 LeuGlyAla 447

RESULT 10
US-10-245-227D-93
; Sequence 93, Application US/10245227D
; Patent No. 6936696
; GENERAL INFORMATION:
; APPLICANT: Monsanto Company
; APPLICANT: Rapp, William D.
; APPLICANT: Peng, Jiexin
; APPLICANT: Nadig, Gautham
; APPLICANT: Venkatesh, Tyamagondlu
; TITLE OF INVENTION: Enhanced Protein and Methods of Their Use
; FILE REFERENCE: REN-00-087 US
; CURRENT APPLICATION NUMBER: US/10/245,227D
; CURRENT FILING DATE: 2002-09-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 93
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Glycine max
US-10-245-227D-93

Alignment Scores:
Pred. No.: 1,29e-104      Length: 449
Score: 1162.00           Matches: 236
Percent Similarity: 70.4%      Conservative: 90
Best Local Similarity: 51.0%    Mismatches: 83
Query Match: 32.4%           Indels: 54
DB: 2                     Gaps: 10

US-10-728-323-1 (1-2032) x US-10-245-227D-93 (1-449)

QY 542 GTGAGGAGAAACATCTCGGAACACCCCTTTCTACTTC---CCGTCAAGGCGGTTTAGC 598
Db 36 ValArgGluAsp-----GluAsnAsnProPheTyrLeuArgSerSerAsnSerPheGln 53
QY 599 ACCCGGTACGGGAACCAAAACGGTAGATCCGGTCTCTCAGAGGTTTGACCAAAAGGTCA 658
Db 54 ThrLeuPheGluAsnGlnAsnGlyArgIleArgLeuGlnArgPheAsnLysArgSer 73
QY 659 AGGCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCAGGCCCAACCTAACACT 718
Db 74 ProGlnLeuGluAsnLeuArgAspTyrArgIleValGlnPheGlnSerLysProAsnThr 93
QY 719 CTTGTTCTCCCAAGCACGCTGATGCTGATAACATCCTTGTATTCCAGCAAGGGCAAGCC 778
Db 94 IleLeuLeuProHisHisAlaAspPheLeuLeuPheValLeuSerGlyArgAla 113
QY 779 ACCGTGACCGTAGCAATGCGAATAACAGAAAGAGCTTTAAATCTTGACGAGGGCCATGCA 838
Db 114 IleLeuThrLeuValAsnAsnAspAspArgAspSerTyrAsnLeuHisProGlyAspAla 133
QY 839 CTCAGAAATCCCATCCGGTTTCATTCTCATCTTGAACCGCCATGACAAACCAAGAACCTC 898
Db 134 GlnArgIleProAlaGlyThrThrTyrLeuValAsnProHisAspHisGlnAsnLeu 153
QY 899 AGAGTAGCTAAATCTCCATGCCGTTTAAACACACCCGGCCAGTTTGAGGATTTCTCCCG 958
Db 154 LysIleIleLysLeuAlaIleProValAsnLysProGlyArgTyrAspAspPhePheLeu 173
QY 959 GCGAGGACGCGACCAATCATCTCTTCGCGGCTTCAGCAGGAATACGTTGGAGGCC 1018
Db 174 SerSerThrGlnAlaGlnSerTyrLeuGlnGlyPheSerHisAsnIleLeuGluThr 193
QY 1019 GCCTTCAATCGGAATTCATAGATACGAGGAGTCTGTTAGAGAGATGACGAGAGGT 1078
Db 194 SerPheHisSerGluPheGluGluIleAsnArgValLeuLeu-----Gly 208
QY 1079 GACCAAGAGAGAGAGGGGACAGCGATGAGTACTCGGAGTAGTGAGAACATGAGGA 1138
Db 209 GluGluGluGluGlnArgGlnGln-----GluGly 218
QY 1139 GTGATAGTCAAAGTGTCAAAGGACGTTGAAGAACTTACTAAGCACGCTAAATCCGTC 1198
Db 219 ValIleValGluLeuSerLysGluGlnIleArgGlnLeuSerArgArgAlaLysSerSer 238
QY 1199 TCAAAGAAA-----GCCTCCGAAGAGAGGGAGATATACCAACCAATCAACTTGAGA 1252
Db 239 SerArgLysThrIleSerSerGluAspGlu-----ProPheAsnLeuArg 253
QY 1253 GAAGGCGAGCCCATCTTTCTCAACATTTGGGAAGTTATTTCAGGTGAGCCAGACAAG 1312
Db 254 SerArgAsnProIleTyrSerAsnAsnPheGlyLysPheGluIleThrProGlu--- 272
QY 1313 AAGAACCCCGAGCTTCAGGACCTGGACATGATGCTACCTGTGTAGAGATCAAAGAGGA 1372
Db 273 LysAsnProGlnLeuArgAspLeuAspIlePheLeuSerSerValAspIleAsnGluGly 292
QY 1373 GCTTTGATGTCCTCCACATTTCAACTCAAAGCCATGTTATCGTCGTCGTCACAAAGGA 1432
Db 293 AlaLeuLeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGly 312
QY 1433 ACTGGAACCTTGAACCTCGGTGTGAAGAAAAGAGCAACAAAGAGGGGACCGCGGAA 1492
Db 313 AspAlaAsnIleGluLeuValGlyIle---LysGluGlnGlnLysGlnLysGlnGlu 331
QY 1493 GAAGAGGAGGACGAAGACGAAGAAGAGGAGGGAAGTAACAGAGAGGTGCGTAGGTACACA 1552
Db 1552
```



```
QY 959 GCGAGAGCGGAGACCAATCATCTACTTCAGGGGCTTCAGCAGGAATACGTTGGAGGCC 1018
Db SerSerThrGlnAlaGlnGlnSerTyLeuGlnGlnGlyPheSerHisAsnIleLeuGluThr 161
QY 1019 GCCTTCAATCGGAATTCAATGATACGAGGGTCTGTGTAGAACAGAAATGCAGGAGGT 1078
Db SerPheHisSerGluPheGluGluIleAsnArgValLeuLeu-----Gly 176
QY 1079 GAGCAAGAGAGAGCGGAGCGGATCGAGTACTCGGAGTAGTGAGAACCAATGAAGGA 1138
Db GluGluGluGluGlnArgGlnGln-----GluGly 186
QY 1139 GTGATAGTCAAAAGTCAAGAGGACAGCTTGAAGAATCTTACTAAGCAGCCTAAATCCGTC 1198
Db ValIleValGluLeuSerLysGluGlnIleArgGlnLeuSerArgAlaLysSerSer 206
QY 1199 TCAAGAGAA-----GGCTCCGAGAGAGGGAGATATCACCAACCAATCAACTTGAGA 1252
Db SerArgLysThrIleSerSerGluAspGlu-----ProPheAsnLeuArg 221
QY 1253 GAAGGGGAGCGGATCTTTCTAAACAATTTGGGAAGTTATTGAGGTGAAGCCAGCACAAG 1312
Db SerArgAsnProIleTySerAsnAsnPheGlyLysPheGluIleThrProGlu--- 240
QY 1313 AAGAACCCCGAGCTTCAGGACCTGGACATGATCTCCTGCTGTAGAGATCAAGAAGGA 1372
Db LysThrProGlnLeuArgAspLeuAspIlePheLeuSerSerValAspIleAsnGluGly 260
QY 1373 GCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCGTCTGCTCAACAAGGA 1432
Db AlaLeuLeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGly 280
QY 1433 ACTGGAACCTTGAACCTCGTGGCTGTAAGAAAAAGCAACAACAGAGGGGACCGCGGAA 1492
Db AspAlaAsnIleGluLeuValGlyIle---LysGluGlnGlnGlnLysGlnLysGlnGlu 299
QY 1493 GAAGAGGAGGACCAAGCAGNAGAGAGGGGAAGTAACAGAGAGGTGCGTAGTACACA 1552
Db GluGluPro-----LeuGluValGlnArgTyArg 309
QY 1553 GCGAGGTGAAGGAGCGATGTGTTTCATCATCCAGCAGCTCATCCAGTAGCCATCAAC 1612
Db AlaGluLeuSerGluAspValPheValIleProAlaIleTyProPheValValAsn 329
QY 1613 GCTTCCTCCGAATCCTCTGCTGGCTTCGGTATCAACGCTCAACCAACCAACCAAGATC 1672
Db AlaThrSerAsnLeuAsnPheLeuAlaPheGlyIleAsnAlaGluAsnGlnArgAsn 349
QY 1673 TTCCTTGACGCTGATTAAGNCAATGTGATAGACAGATAGAGAAGCAAGCGAAGGATTTA 1732
Db PheLeuAlaGlyGluLysAspAsnValValArgGlnIleGluArgGlnValGlnGluLeu 369
QY 1733 GCATTCCCTCGGTGCGGTGAACAAGTTGAGAAGCTCATCAAAAACCAAGAGGAATCTCAC 1792
Db AlaPheProGlySerAlaGlnAspValGluArgLeuLeuLysLysGlnArgGluSerTy 389
QY 1793 TTTGTGAGTCTCGTCTCTCAATCTCAATCTCAATCTCCGTCGTCTCCTGAGAAGAGTCT 1852
Db PheValAspAlaGlnProGln----- 396
QY 1853 CCTGAGNAAGAGATCAAGAGGAGGAAACCAAGGAGGGAGGTCCTCACTCTTTCAATT 1912
Db -----GlnLysGluGluGlySerLysGlyArgLysGlyProPheProSerIle 412
QY 1913 TTGAAGGCT 1921
Db 413 LeuGlyAla 415
```

RESULT 12

US-10-245-227D-12

; Sequence 12, Application US/10245227D

; Patent No. 6936696

```
; GENERAL INFORMATION:
; APPLICANT: Monsanto Company
; APPLICANT: Rapp, William D.
; APPLICANT: Peng, JieXin
; APPLICANT: NadiG, Gautham
; APPLICANT: Venkatesh, Tyamagondlu
; TITLE OF INVENTION: Enhanced Protein and Methods of Their Use
; FILE REFERENCE: REN-00-087 US
; CURRENT APPLICATION NUMBER: US/10/245,227D
; CURRENT FILING DATE: 2002-09-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 12
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: epitope form of beta-conglycinin beta subunit
US-10-245-227D-12
```

```
Alignment Scores:
Pred. No.: 3.83e-104 Length: 425
Score: 1157.00 Matches: 237
Percent Similarity: 70.0% Conservative: 87
Best Local Similarity: 51.2% Mismatches: 85
Query Match: 32.3% Indels: 54
DB: 2 Gaps: 10
```

US-10-728-323-1 (1-2032) x US-10-245-227D-12 (1-425)

```
QY 542 GTGAGGAAGAAACATCTCGGAACACCTTTCTACTTC--CCGTCAAGCGGTTTAGC 598
Db 12 ValArgGluLeu-----GluAsnAsnProPheTyLeuArgSerSerAsnSerPheGln 29
QY 599 ACCCGCTAGCGGAACCAAAACCGTAGGATCCGGGCTCTCGAGAGGTTTGACCAAGGTCA 658
Db 30 ThrLeuPheGluAsnGlnAsnGlyArgIleArgLeuLeuGlnArgPheAsnLysArgSer 49
QY 659 AGCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAACCTAACACT 718
Db 50 ProGlnLeuGluAsnLeuArgAspTyArgIleValGlnPheGlnSerLysProAsnThr 69
QY 719 CTGTGTTCTCCAAAGCACCGCTGATGCTATAACATCCTTGTATTCCAGCAAGGGCAAGCC 778
Db 70 IleLeuLeuProHisIleAlaAspAlaAspPheLeuLeuPheValLeuSerGlyArgAla 89
QY 779 ACGCTGACGCTAGCAAAATGGCAATAACAGAAAAGAGCTTTAATCTTCAGAGGCCCATGCA 838
Db 90 IleLeuThrLeuValAsnAsnAspArgAspSerTyAsnLeuHisProGlyAspAla 109
QY 839 CTCAGAAATCCATCCCGTTTCATTTCTTACATCTTGAACCGCCATCACACACCAAGACCTC 898
Db 110 GlnArgIleProAlaGlyThrThrTyTyLeuValAsnProHisAspHisGlnAsnLeu 129
QY 899 AGAGTAGCTAAATCTCCATGCCCGTTAAACACCCCGCCAGTTTCAGGATTTCTTCCCG 958
Db 130 LysIleIleLysLeuAlaIleProValAsnLysProGlyArgTyArgAspPhePheLeu 149
QY 959 GCGAGCAGCGGAGACCAATCATCTACTTCGAGGGCTTCAGCAGGAATACGTTGGAGGCC 1018
Db 150 SerSerThrGlnAlaGlnGlnSerTyLeuGlnGlnGlyPheSerHisAsnIleLeuGluThr 169
QY 1019 GCCTTCAATCGGAATTCATGAGATACGAGGGTCTGTTAGAAGAGATGCGAGGAGGT 1078
Db 170 SerPheHisSerGluPheGluGluIleAsnArgValLeuLeu-----Gly 184
QY 1079 GAGCAAGAGAGAGAGGGCAGAGCGGATCGAGTACTCGGAGTAGTGAGAACCAATGAAGGA 1138
Db 185 GluGluGluGluGlnArgGlnGln-----GluGly 194
QY 1139 GTGATAGTCAAAAGTCAAGAGGAGCAGCTTGAAGAATCTTACTAAGCAGCCTAAATCCGTC 1198
Db 195 ValIleValGluLeuSerLysGluGlnIleArgGlnLeuSerArgAlaLysSerSer 214
```

QY 1199 TCAAGAAA-----GGCTCCGAAGAGAGGAGATATCACCAACCCCAATCACTTGAGA 1252
DB 215 SerArgLysThrIleSerSerGluAspGlu-----ProPheAsnLeuArg 229
QY 1253 GAAGGGAGCCGATCTTCTTCAACACTTTGGGAAGTATTGAGGTGAAGCAGACAAAG 1312
DB 230 SerArgAsnProIleTyrSerAsnAsnPheGlyLysPheGluIleThrProGlu--- 248
QY 1313 AAGAACCCTCAGCTCAGACCTGGACATGATGCTCACCTGCTGAGATCAAGAAGGA 1372
DB 249 LysThrProGlnLeuArgAspLeuAspIlePheLeuSerSerValAspIleAsnGluGly 268
QY 1373 GCTTTGATCTCCACACTTCAACTCAAGGGCATTGGTTATCGTCTGCTCAACAAGGA 1432
DB 269 AlaLeuLeuLeuProHisPheAsnSerLysAlaIleValIleuValIleAsnGluGly 288
QY 1433 ACTGAAACCTTGAACTCTGCTGTGAAGAAAGAGCAACAACAGAGGGGAGCGGGGAA 1492
DB 289 AspAlaAsnIleGluLeuValGlyIle---LysGluGlnGlnLysGlnLysGlnGlu 307
QY 1493 GAAGAGGAGGAGCAAGACGAAGAGGAGGAGTAAACAGAGAGGTGCTAGGTACACA 1552
DB 308 GluGluPro-----LeuGluValGlnArgTyrArg 317
QY 1553 GCGAGGTTGAAGAGGCGATGTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAAC 1612
DB 318 AlaGluLeuSerGluAspValPheValIleProAlaIleTyrProPheValValAsn 337
QY 1613 GCTTCTCCGAATCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1672
DB 338 AlaThrSerAsnLeuAsnPheLeuAlaPheGlyIleAsnAlaGluAsnAsnGlnArgAsn 357
QY 1673 TTCTTGCAGGTGATGAAGACATGTGATGACGATAGACGAGCAAGCAAGCAAGTCTCA 1732
DB 358 PheLeuAlaGlyLysAspAsnValValArgGlnIleGluArgGlnValGlnGluLeu 377
QY 1733 GCATTCCTCGGTGCGGTGCAACAAAGTTGAGAAAGCTCATCAAAAACAGAAAGCAATCTCA 1792
DB 378 AlaPheProGlySerAlaGlnAspValGluArgLeuLeuLysLysGlnArgGluSerTyr 397
QY 1793 TTTGTGAGTGTCTCTCAATCTCAATCTCAATCTCAATCTCTCTCTCTCTCTCTCTCT 1852
DB 398 PheValAspAlaGlnProGln----- 404
QY 1853 CTGAGAAGAGGATCAAGAGGAGGAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1912
DB 405 -----GlnLysGluGluGlySerLysGlyArgLysGlyProPheProSerIle 420
QY 1913 TTGAAGGCT 1921
DB 421 LeuGlyAla 423

RESULT 13

US-10-245-227D-2

; Sequence 2, Application US/10245227D

; Patent No. 6936696

; GENERAL INFORMATION:

; APPLICANT: Monsanto Company

; APPLICANT: Rapp, William D.

; APPLICANT: Peng, Jixen

; APPLICANT: Nadiq, Gautham

; APPLICANT: Venkatesh, Tyamagondlu

; TITLE OF INVENTION: Enhanced Protein and Methods of Their Use

; FILE REFERENCE: REN-00-087 US

; CURRENT APPLICATION NUMBER: US/10/245, 227D

; CURRENT FILING DATE: 2002-09-17

; NUMBER OF SEQ ID NOS: 97

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 2

; LENGTH: 390

; TYPE: PRT

; ORGANISM: Glycine max

US-10-245-227D-2

Alignment Scores:

Pred. No.: 2,67e-100 Length: 390
Score: 1117.50 Matches: 225
Percent Similarity: 72.6% Conservative: 83
Best Local Similarity: 53.1% Mismatches: 79
Query Match: 31.2% Indels: 37
DB: 2 Gaps: 9

US-10-728-323-1 (1-2032) x US-10-245-227D-2 (1-390)

QY 542 GTGAGGGAAGAAACATCTCGGAACAAACCCCTTTCTACTTC---CCGTCAGAGCGGTTTACG 598
DB 1 ValArgGluAsp-----GluAsnAsnProPheTyrPheArgSerSerAsnSerPheGln 18
QY 599 ACCCGCTACGGGAACCAAAACGGTAGGATCCGGTCTCGCAGAGGCTTTCACCAAAAGGTCA 658
DB 19 ThrLeuPheGluAsnGlnAsnValArgIleArgLeuLeuGlnArgPheAsnLysArgSer 38
QY 659 AGGCAATTTCCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTAACACT 718
DB 39 ProGlnLeuGluAsnLeuArgAspTyrArgIleValGlnPheGlnSerLysProAsnThr 58
QY 719 CTTGTTCTCCCAAGCAGCTGATGCTGATACATCTTGTATTATCCAGCAAGGGCAAGCC 778
DB 59 IleLeuLeuProHisHisAlaAspAlaAspPheLeuLeuPheValLeuSerGlyArgAla 78
QY 779 ACCGTGACGTAGCAAAATGGCAATTAACAGAAAGAGCTTTAATCTTTCAGCAGGGGCCATGCA 838
DB 79 IleLeuThrLeuValAsnAsnAspArgAspSerTyrAsnLeuHisProGlyAspAla 98
QY 839 CTCAGAAATCCCATCCCGTTTCATTCTCTACATCTTGAACCGCCATCAACCAAGAACCTTC 898
DB 99 GlnArgIleProAlaGlyThrThrTyrTyrLeuValAsnProHisAspHisGlnAsnLeu 118
QY 899 AGAGTAGTAAATCTCCATGCCCTTAAACACACCCGCCAGCTTTGAGAGATTTCTTCCCG 958
DB 119 LysIleIleLysLeuAlaIleProValAsnLysProGlyArgTyrAspAspPhePheLeu 138
QY 959 GCGACAGCCGAGACCAATCATCTCTGAGGGCTTCAGCAGGAATACGTTGGAGGCC 1018
DB 139 SerSerThrGlnAlaGlnGlnSerTyrLeuGlnGlyPheSerHisAsnIleLeuGluThr 158
QY 1019 GCCTTCAATCGGAATTCATAGATACGAGGGTCTGTTAGAGAAGAGAATCAGAGAGT 1078
DB 159 SerPheHisSerGluPheGluGluIleAsnArgValLeuPhe-----Gly 173
QY 1079 GAGCAAGAGGAGAGGAGGAGGAGGAGGAGTACGAGTACTCGGAGTAGTGAGAACATGAAGGA 1138
DB 174 GluGluGluGluGlnArgGlnGln-----GluGly 183
QY 1139 GTGATAGTCAAGAGTGTCAAGGAGCAGCTTGAAGAACTTACTTAAGCACGCTAAATCCGTC 1198
DB 184 ValIleValGluLeuSerLysGluGlnIleArgGlnLeuSerArgArgAlaLysSerSer 203
QY 1199 TCAAAAGAAA-----GGCTCCGAAGAGAGGAGAGATATCACCAACCCCAATCACTTGAGA 1252
DB 204 SerArgLysThrIleSerSerGluAspGlu-----ProPheAsnLeuArg 218
QY 1253 GAAGCGAGCCGATCTTTTCTAAACACTTTGGGAAGTATTATTGAGGTGAAGCAGACAAAG 1312
DB 219 SerArgAsnProIleTyrSerAsnAsnPheGlyLysPheGluIleThrProGlu--- 237
QY 1313 AAGAACCCTCAGCTCAGACCTGGACATGATGCTCACCTGCTGAGATCAAGAAGGA 1372
DB 238 LysAsnProGlnLeuArgAspLeuAspIlePheLeuSerSerValAspIleAsnGluGly 257
QY 1373 GCTTTGATCTCCACACTTCAACTCAAGGGCATTGGTTATCGTCTGCTCAACAAGGA 1432
DB 258 AlaLeuLeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGly 277
QY 1433 ACTGAAACCTTTGAACTCGTCTGTAGAAAAGAGCAACAACAGAGGGGAGCGGGGAA 1492

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 23:41:32 ; Search time 6.9905 Seconds
(without alignments)
4094.130 Million cell updates/sec

Title: US-10-728-323-1
Perfect score: 3586
Sequence: 1 aataatcatatatattcatc.....cgttgtgctgtttcttccc 2032

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 250354 seqs, 46948837 residues

Total number of hits satisfying chosen parameters: 500708

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-O=/abs/ABSSWEB.spool/US10728323/runat.15052006.172148.22650/app.query.fasta.1
-DB=PublishedApplications_AA_New -QFMT=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abs02p
-USER=US10728323 @CN 1.1.63 @runat.15052006.172148.22650 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :				Published Applications AA New.*	
				1: /SID55/ptodata/1/pubpaa/US08_NEW_PUB.pep.1*	
				2: /SID55/ptodata/1/pubpaa/US06_NEW_PUB.pep.1*	
				3: /SID55/ptodata/1/pubpaa/US07_NEW_PUB.pep.1*	
				4: /SID55/ptodata/1/pubpaa/US08_NEW_PUB.pep.1*	
				5: /SID55/ptodata/1/pubpaa/PCT_NEW_PUB.pep.1*	
				6: /SID55/ptodata/1/pubpaa/US09_NEW_PUB.pep.1*	
				7: /SID55/ptodata/1/pubpaa/US09_NEW_PUB.pep.1*	
				8: /SID55/ptodata/1/pubpaa/US10_NEW_PUB.pep.1*	
				9: /SID55/ptodata/1/pubpaa/US10_NEW_PUB.pep.1*	
				10: /SID55/ptodata/1/pubpaa/US11_NEW_PUB.pep.1*	
				11: /SID55/ptodata/1/pubpaa/US11_NEW_PUB.pep.1*	
				12: /SID55/ptodata/1/pubpaa/US60_NEW_PUB.pep.1*	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	ID Description
1	3286	91.6	626 9 US-10-498-026-89 Sequence 89, Appl
2	3286	91.6	626 11 US-11-033-039-10 Sequence 10, Appl
3	228.5	6.4	499 9 US-10-508-263-94 Sequence 94, Appl
4	179	5.0	8746 11 US-11-098-686-10232 Sequence 10232, A
5	172	4.8	798 11 US-11-072-512-3335 Sequence 3335, Ap

6	170.5	4.8	507	11	US-11-033-039-96	Sequence 96, Appl
7	166.5	4.6	481	9	US-10-508-263-28	Sequence 28, Appl
8	163.5	4.6	651	11	US-11-128-660-1	Sequence 1, Appl
9	162.5	4.5	499	11	US-11-087-099-1159	Sequence 1159, Ap
10	162	4.5	5689	9	US-10-784-004-709	Sequence 709, App
11	157	4.4	495	9	US-10-508-263-20	Sequence 20, Appl
12	153	4.3	677	9	US-10-982-545-12	Sequence 12, Appl
13	149	4.2	623	11	US-11-072-512-3378	Sequence 3378, Ap
14	148	4.1	562	9	US-10-508-263-24	Sequence 24, Appl
15	147.5	4.1	715	11	US-11-072-512-3385	Sequence 3385, Ap
16	146	4.1	1719	11	US-11-234-786-378	Sequence 378, App
17	145.5	4.1	656	9	US-10-784-004-451	Sequence 451, App
18	145	4.0	1197	11	US-11-087-099-8238	Sequence 8238, Ap
19	144	4.0	1212	9	US-10-979-095-9	Sequence 9, Appl
20	143	4.0	671	11	US-11-234-786-380	Sequence 380, App
21	143	4.0	671	11	US-11-139-041-306	Sequence 306, App
22	142.5	4.0	656	11	US-11-234-786-379	Sequence 379, App
23	142.5	4.0	656	11	US-11-139-041-305	Sequence 305, App
24	140.5	3.9	2910	9	US-10-330-773-39	Sequence 39, Appl
25	139	3.9	2004	9	US-10-469-469-250	Sequence 250, App
26	139	3.9	4373	10	US-11-118-524-2	Sequence 2, Appl
27	139	3.9	4374	11	US-11-128-572-2	Sequence 2, Appl
28	136.5	3.8	2426	11	US-11-203-806A-11	Sequence 11, Appl
29	135.5	3.8	2472	11	US-11-203-806A-2	Sequence 2, Appl
30	134	3.7	585	9	US-10-784-004-712	Sequence 712, App
31	134	3.7	795	9	US-10-770-726-49	Sequence 49, Appl
32	133.5	3.7	594	9	US-10-131-826A-10	Sequence 10, Appl
33	133.5	3.7	594	9	US-10-973-115B-10	Sequence 10, Appl
34	133.5	3.7	594	9	US-10-137-873A-10	Sequence 10, Appl
35	133.5	3.7	594	9	US-10-152-370-10	Sequence 10, Appl
36	133.5	3.7	594	11	US-11-290-153-10	Sequence 10, Appl
37	132.5	3.7	676	11	US-11-135-855-28	Sequence 28, Appl
38	132.5	3.7	717	11	US-11-135-855-29	Sequence 29, Appl
39	132.5	3.7	841	11	US-11-294-997-37	Sequence 37, Appl
40	131.5	3.7	2344	9	US-10-330-773-627	Sequence 627, App
41	131	3.7	671	8	US-10-196-749-346	Sequence 346, App
42	131	3.7	671	9	US-10-194-487-346	Sequence 346, App
43	131	3.7	671	9	US-10-195-883-346	Sequence 346, App
44	131	3.7	671	9	US-10-195-888-346	Sequence 346, App
45	131	3.7	671	9	US-10-195-889-346	Sequence 346, App

ALIGNMENTS

RESULT 1
US-10-498-026-89
; Sequence 89, Application US/10498026
; Publication No. US20060024334A1
; GENERAL INFORMATION:
; APPLICANT: CIRCASSIA LIMITED
; TITLE OF INVENTION: IMMUNOTHERAPEUTIC METHODS AND SYSTEMS
; FILE REFERENCE: N.87430 WO GCW
; CURRENT APPLICATION NUMBER: US/10/498,026
; CURRENT FILING DATE: 2004-06-04
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 89
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-498-026-89

Alignment Scores:
Pred. No.: 3,43e-262 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.6% Indels: 0
DB: 9 Gaps: 0

US-10-728-323-1 (1-2032) x US-10-498-026-89 (1-626)

QY 50 ATGAGAGGAGGGTTCTTCCTCACTGATGCTGTGCTAGGATCCTTGCTCTGCTTCAGTT 109

|||||
1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGATGCCCAAGTATCATCACCTTACCAAGAAAGAAACAGAGAACCCCTGCGCCAG 169
Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGTGCCTCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCT 229
Db 41 ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGCTGCACCAAGCTCCAGATATGATCCTCGTTGTCTATGATCCTCGAGACACACTGCG 289
Db 61 ArgCysThrLysLeuGlnLysArgProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAACGTTCCCTCCAGGGAGCGGACACGCTGGCGGCCAACCCGGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCGCTCAACCCGAGAGAGAGAGAGGCGCGATGGGACACGCTGGACCG 409
Db 101 AspAspAspArgArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGlyPro 120
QY 410 AGGGAGCGTCAAGAGAAAGAACTCGAGACCAAGAGAGAGATTGGAGGCGACCAAGT 469
Db 121 ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgProSer 140
QY 470 CATCAGCAGCCCGGAAAAATAAGCCCGAAGGAGAGAGAGAAACAGAGTGGGAAACA 529
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGlnTrpGlyThr 160
QY 530 CAGGTTAGCCATGTGAGGGAAGAAACATCTCGGAACAAACCTTTCTACTTCCGCTCAAG 589
Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTTAGCACCGCTACGGGAACCAAAACGGTAGGATCCGGTCTCGCAGAGGTTTGAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGTCAAGGCAGTTTCAGAACTCTCCAGAAATCCACCGTATTGTGCAGATCGAGGCCAAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTTTGTTCTTCCAGCAGCGTGATGCTGATAACATCTTGTATTATCCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTAGCCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAG 829
Db 241 GlyGlnAlaThrValThrValAlaAsnGlnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCCATGCATCAGAAATCCCATCCGTTTCATTTCCCTACATCTTGAAACCGCCCATGACAAC 889
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACTCAGACTAGCTAAATCTCCATCCCGCTTAACACACCCGCCAGTTTGAGGAT 949
Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCGAGCAGCCGAGACCAATCATCTACTTGCAGGCGCTTCAGCAGGAATACG 1009
Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGAGGCGCCCTTCAATGCGGAATTCATATGATACGAGGAGTGTGTTAGAGAGAAT 1069
Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGGTGACCAAGGAGAGAGGCGCAGCGGATGGAGTACTCGGACTACTGAGAAC 1129
Db 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGGAGTGTAGTCAAAGTGTCAAAGGACGCTTGAAGAACTTACTAAGCAGCGT 1189
|||||

Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AAATCCGTCTCAAAGAAAGGCTCCAAAGAGAGGAGATATCACCAACCAATCACTTG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 ACAGAAAGCGGACCGCATCTTCTTAACAACATTTGGGAAGTATTATTGAGGTGAAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAAACCCCGACTTCAGGACCTCGGACATGATCTCACCTGTGTAGAGATCAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTGTATGCTCCACACTTCAACTCAAAAGGCCATGGTTATCGTCTGCTCAACAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
QY 1430 GGAACCTGGAACCTTGAACCTCGTGTAAAGAAAGAACAAACAGAGGGGACGCGG 1489
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnArgGlyArgArg 480
QY 1490 GAAGAAGAGGAGCAGCAGAACGAAAGAGAGGGAAGTAACAGAGAGGTGCGTAGGTAC 1549
Db 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGGAGGTTGAAGGAGGCGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AAGCTTCTCTCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTCAAAACCAACCAACA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTCTTCAGGTGATAAGGACAAATGTGATAGACCATAGAGAGAGCAAGCAAGGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCAATTCCTGGTGGTGGTGAACAAGTTGAGAAGCTCATCAAAACCAAGAGGATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCCGTCGTCTCTGAGAAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTGAGAAAGAGATCAAGAGGAGGAGAAACCAAGGAGGAGGAGGTCCTCTCTTCA 1909
Db 601 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGCTTTTAAAC 1927
Db 621 IleLeuLysAlaPheAsn 626
RESULT 2
US-11-033-039-10
; Sequence 10, Application US/11033039
; Publication No. US20060002947A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: LI-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2017US01
; CURRENT APPLICATION NUMBER: US/11/033,039
; PRIOR FILING DATE: 2005-01-11
; PRIOR APPLICATION NUMBER: 10/245,871
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 1452
; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 10
 ; LENGTH: 626
 ; TYPE: PRT
 ; ORGANISM: Arachis hypogaea
 ; US-11-033-039-10

Alignment Scores:

Pred. No.: 3,43e-262 Length: 626
 Score: 3286.00 Matches: 626
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 91.6% Indels: 0
 DB: 11 Gaps: 0

US-10-728-323-1 (1-2032) x US-11-033-039-10 (1-626)

QY 50 ATGAGAGGAGGTTTCTCCACTGATGCTTGCTAGGGATCCTTGCTCGCTTCAGTT 109
 Db 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyLeuValLeuAlaSerVal 20
 QY 110 TCTGCAACGCATGCCAAGTCATCACCTTACAGAGAGAAACAGAGAACCCCTGCGCCAG 169
 Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
 QY 170 AGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCT 229
 Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
 QY 230 CCCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGGC 289
 Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
 QY 290 ACCACCAACCAAGTTCCCTCCAGGGAGCGGACACGTCGGCGCCCAACCCGGAGACTAC 349
 Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
 QY 350 GATGATACCGCGCTCAACCCCGAAGAGAGAGAGGCGCGATGGGACCACTGACCG 409
 Db 101 AspAspAspArgArgGlnProArgArgGluGluGlyGlyArgTyrGlyProAlaGlyPro 120
 QY 410 AGGGAGCGTGAAGAGAGAACTGGAGACACCAACAGAGAGATTTGGAGGCGACCAAGT 469
 Db 121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
 QY 470 CATCAGCACCGAAGAGAAATAGGCCGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
 Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
 QY 530 CCAGTATGCTGTGAGGAGAAACATCTCGGAACAACCCCTTCTACTTCCCGTCAAGG 589
 Db 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
 QY 590 CGTTTATGACCGCTACGGGAACCAACCGTAGATCCGGGTCTCGAGAGTTTGAC 649
 Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
 QY 650 CAAAGTCAAGCAGTTTCAGAAATCTCCAGAAATCCCGTATTGTGCAGATCGAGGCCAAA 709
 Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
 QY 710 CCTAACACTCTTGTCTTCCCAAGCACGCTGATGCTGATAACATCTTGTATCCAGCAA 769
 Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
 QY 770 GGGCAAGCCACGCTGACCGTAGCAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAG 829
 Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
 QY 830 GSCCATGCACTCAGAAATCCATCCGGTTTCATTTCTCATCTTGAACCGCCATCACAC 889
 Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
 QY 890 CAGAACCTCAGAGTAGTAAATCTCCATGCCGTTAAACACACCCGCGCAGTTTGAGGAT 949

Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
 QY 950 TTCCTCCCGCGAGCAGCGAGACCAATCATCTCTACTTCGAGGCTTCAGCAGGAATAGC 1009
 Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
 QY 1010 TTGAGAGCGCCCTTCAATGCGGAATTCAATGAGATACGAGGCTGCTGTAGAGAGAAT 1069
 Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 340
 QY 1070 GCAGAGGTGAGCAAGAGAGAGAGCGGAGGCGATGGAGTACTCGGAGTAGTGAGAAC 1129
 Db 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTyrPheSerThrArgSerSerGluAsn 360
 QY 1130 AATGAAGAGATGATAGTCAAGAGTCAAGAGGACGCTGAGAACTTACTTAAGCAGCT 1189
 Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
 QY 1190 AAATCCGTCTCAAAGAAAGGCTCCCAAGAGAGGAGATATCACCAACCCCAATCAACTTG 1249
 Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
 QY 1250 AGAGAAAGCGAGCCCGATCTTTTAACTTTGGAAAGTTATTGAGGTGAAGCCAGAC 1309
 Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
 QY 1310 AAGAGAACCCCGCTTCAGGACCTGCAGCATGATGCTCACCTGTGTAGAGATCAAGAA 1369
 Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
 QY 1370 GGAGTTTGATGCTCCACACTTCAACTCAAAGGCGCATGTTATCGTCTCGTCAACAAA 1429
 Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
 QY 1430 GGAACCTGAAACCTTGAACTCGTGGCTTAAGAAAAGCAACAACAGAGGGAGCGCGG 1489
 Db 461 GlyThrGlyAsnLeuLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
 QY 1490 GAAGAGAGGAGGAG 1549
 Db 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
 QY 1550 ACAGCGAGTTGAAGAAAGCGCATGTTTCATCATGCCAGAGCTCATCCAGTAGCCATC 1609
 Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
 QY 1610 AACGCTTCCCTCCGAACCTCGTCTCGCTTCGCTTCAACGCTGAAACCAACCAACAGA 1669
 Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
 QY 1670 ATCTTCTTCAGGTGATAAGGACCAATGTGATAGACCATAGACAGACAGAGAGAGAT 1729
 Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
 QY 1730 TTAGCATTTCCCTGGGTCCGGTGAACAAGTTGAGAAAGCTCATCAAAAACAGAGAAATCT 1789
 Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
 QY 1790 CACTTTGTGAGTGTCTGCTCCTCAATCTCAATCTCAATCTCCGTCTCTCTGAGAAAGAG 1849
 Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
 QY 1850 TCTCTGAGAAAGAGGATCAAGAGGAGAAACCAAGGAGGAGGAGGCTCCACTCTTCA 1909
 Db 601 SerProGluLysGluAspGlnGluGlnGlnGlyLysGlyLysGlyProLeuLeuSer 620
 QY 1910 ATTTTGAAGGCTTTTAAAC 1927
 Db 621 IleLeuLysAlaPheAsn 626

RESULT 3

US-10-508-263-94

Db	553	lulysGlu-----ArgAspArgGluArgAspA	562
QY	1376	TTGATGCTCCACACTTCAACTCAAGGCCATGGTTATCGTCGTCTCAACAAAGCAACT	1435
Db	562	rgAsp-----ArgAspHisAspArgGluArgGluArgGlu-	575
QY	1436	GGAAACCTTGAACCTCGTGCTGTGAAGAAAGAGCAACAACAGAGGGGACGGCGGGAAGAA	1495
Db	576	ArgAsp-----ArgGluLysGluArgGluArgGluArgGlu	588
QY	1496	GAGGAGGACGAACAGCAAGAGAGGAGGAAGTAACAGAGAGGTGCGTGGTACACAGCG	1555
Db	589	ArgGluArgGluArgGluArgGluArgGluArgGluArgGluArgGluArgGlu	608
QY	1556	AGTTCAAGGAAGGCCATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCT	1615
Db	609	ArgAlaArgGluArgAsp	614
QY	1616	TCTCCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACACAGAATCTTC	1675
Db	615	-----LysGluArgGluArgGlnArg	621
QY	1676	CTTGCAAGTGATAGACAAATGTGATAGACCAGATAGAGAACAGCGAAGATTATA	1732
Db	622	AspTrpGluAspLysAspLysGlyArgAspAspArgArgGluLysArgGluLysArg	641
QY	1733	-----GCATTCCTCGGTGGTGAAACAAGTTGAGAAGCTCATCAAAAACCAAGAGAA	1786
Db	642	GluAspArgAsnProArgAspGlyHisAspGluArgLysSerLysLysArgTyrArgAsn	661
QY	1787	TCTCACTTTGTGAGTGCTCGTCTCAATCTCAATCTCAATCTCCGTCGCTCTGAGAAA	1846
Db	662	GluGlySerProSerProArgGlnSerProLysArgArgGluHisSerProAspSer	681
QY	1847	GAGTCTCTCAGNAAGAGATCAAGAGGAGGAGAAACCAAGGAGGAGGTCCATCTCTT	1906
Db	682	AspAlaTyrAsnSerGlyAspLysAsnGlu-----LysHisArgLeuLeu	697
QY	1907	TCA 1909	
Db	698	Ser 698	
RESULT 6			
US-11-033-039-96			
; Sequence 96, Application US/11033039			
; Publication No. US2006002947A1			
; GENERAL INFORMATION:			
; APPLICANT: HUMPHREYS, ROBERT			
; TITLE OF INVENTION: LI-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES			
; FILE REFERENCE: REH-2017US01			
; CURRENT APPLICATION NUMBER: US/11/033,039			
; CURRENT FILING DATE: 2005-01-11			
; PRIOR APPLICATION NUMBER: 10/245,871			
; PRIOR FILING DATE: 2002-09-17			
; PRIOR APPLICATION NUMBER: 10/197,000			
; PRIOR FILING DATE: 2002-07-17			
; PRIOR APPLICATION NUMBER: 09/396,813			
; PRIOR FILING DATE: 1999-09-14			
; NUMBER OF SEQ ID NOS: 1452			
; SOFTWARE: Patent in version 3.3			
; SEQ ID NO 96			
; LENGTH: 507			
; TYPE: PRT			
; ORGANISM: Arachis hypogaea			
US-11-033-039-96			
Alignment Scores:			
Pred. No.:	8,85e-06	Length:	507
Score:	170.50	Matches:	116
Percent Similarity:	34.5%	Conservative:	77
Best Local Similarity:	20.7%	Mismatches:	214
Query Match:	4.8%	Indels:	153

DB:	11	Gaps:	23
US-10-728-323-1 (1-2032) x US-11-033-039-96 (1-507)			
QY	437	AGACAAACCAAGAGAGAT-----TGGAGGCCAACCAAGTCATCAGCAGCCA-----CGG	484
Db	2	GlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAspAsn	21
QY	485	AAAATAAGGCCCGCAAGGA-----AGAGAAGGAGAACCAAGCTGGCGA	526
Db	22	ArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPheGlu	41
QY	527	ACACAGGTAGCATGTGAGGAGAAACATCTCGAAACAAC-----CCTTTC	574
Db	42	CysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgProPhe	61
QY	575	TAC-----TTCCCGTCAAGGCGGTTTAGCACCCTAGCAGGAAACCAACCGTAGGATC	628
Db	62	TyrSerAsnAlaProGlnGluIlePheIle-----GlnGlnGlyArgGly	76
QY	629	CGGTCCTGCAGAGGTTTGACCAAGGTCAGGCAGTTTTCAGAAATCTCCAGAATCAC---	685
Db	77	TyrPheGlyLeuIlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGly	96
QY	686	CGTATTGTGCAGATCGAGGCCAAACCTAACACTCTTGTCTTCTTCCCAAGCACGCTGATGT	745
Db	97	ArgArgSerGlnSerGlnArgProArgArgLeu-----	108
QY	746	GATAACATCTTGTATCCAGCAAGGCAAGCCCGTGACCGTAGCAATCGCAATAAC	805
Db	109	-----GlnGlyGluAspGlnSerGlnGlnGlnArgAspSerHis	121
QY	806	AGAAAGAGCTTTAATCTTGACGAGGCGCATGCACTCAGAAATCCCATCCGTTTTCATTTCC	865
Db	122	GlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAlaPhe	141
QY	866	TACATCTTGAACCGCCATGACAAACCAAGCCTCAGAGTAGCT-----	907
Db	142	TrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsnAsn	161
QY	908	-----AAAATCTCATGCTCCGTTTAAACACACACCCGCG	937
Db	162	AsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGluGln	181
QY	938	CAGTTTGAGGATTTCTTCCCGCGCAGC-----AGCCGAGACCAATCATCTTACTTGCAG	991
Db	182	GluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyrSer	201
QY	992	GGCTTCAGCAGGAATACGTTGGAGCCGCTTCAATGCGGAATTCATAGATACGAGG	1051
Db	202	ProTyrSerProGlnSerGlnProArgGlnGluArgGluPheSerProArgGlyGln	221
QY	1052	GTGCTGTTAGAGAGAATTCAGGAGGTGAGCAAGAGGAGAGAGG-----	1096
Db	222	HisSerArgArgGluArgAlaGlyGlnGluGluGluAsnGluGlyAsnIlePheSer	241
QY	1096	-----	1096
Db	242	GlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleValGln	261
QY	1097	CAGAGCGCATGAGTACTCGGAGTAGTGAGAAACAATGAAGGAGTGATAGTCAAAGTG---	1153
Db	262	AsnLeuArgGlyGluThrGluSerGluGlu-----GluGlyAlaIleValThrValArg	279
QY	1154	-----TCAAAGGAGCAC	1165
Db	280	GlyGlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluTyr	299
QY	1166	GTTGAAGAACTTACTAAGCACGCTAAATCCGCTCAAAGAAAGGCTCCCAAGAAAGAGGA	1225
Db	300	AspGluAspGluTyrGluTyrAspGluGluAspArgArgArgGlyArgGlySerArgGly	319
QY	1226	-----GATATCACCACCAATCAACTTTGAGAGAGGCGAGCCGATCTTTCTAAC	1276


```
QY 1685 GATAAGGACAAAT-----GTGATAGACCAGATAGAGAACGCG 1723
|||||:|||||
Db 519 AspLysGluAsnIleSerLysGluAsnAspValLeuAspGluLysGluGluAla 538
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1724 AAGGATTAGCATTCCTGGGTGGGTGAACAAAGTTGAGAAGCTCATCAAAAAACCAAGAAG 1783
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 539 GluGluThr-----GluGluGluGluLeuGluGluLysAsnGluGlu 552
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1784 GAATCTCACTTTGTGAGTGTCTCTCAATCTCAATCTCAATCTCAATCTCGTCTCGTAG 1843
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 553 Glu-----ThrGluSerGluIleSerGluAsp 561
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1844 AAAGAGTCTCTGAGAAAGAGGATCAAGAGGAGGAAACCAAGAGGGAAG 1894
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 562 GluGluGluGluGluGluGluGluGluGluGluGluGluGluGluLysLys 578
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 9
US-11-087-099-1159
; Sequence 1159, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 1159
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-11-087-099-1159

Alignment Scores:
Pred. No.: 4,01e-05 Length: 499
Score: 162.50 Matches: 54
Percent Similarity: 45.7% Conservative: 51
Best Local Similarity: 23.5% Mismatches: 81
Query Match: 4.5% Indels: 44
DB: 11 Gaps: 6

US-10-728-323-1 (1-2032) x US-11-087-099-1159 (1-499)
QY 1252 AGAAGCGACCCGATCTTTCTAACACTTTGGGAGTTATTTGAGTGAGCCAGAC-- 1309
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 57 ArgArgAsnLysLysLysSerSerGlnLeuPheAsnVal-PheGlnGluLysProAsp 76
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1310 -----AAGAGAAACCCCGAGCTTCAGGA 1332
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 76 eGluAsnCysAsnGlyTrpSerThrValIleAsnArgLysLeuProAlaLeuLysG 96
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1333 CTGGACATGATGCTACCTGTGTAGATCAAGAGGAGCTTTGATGCTCCACACTT 1392
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 96 ySerGlnIleGlyIleTyrValValAsnLeuThrLysGlySerMetMetGlyProHis 116
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1393 CAACTCAAGCCATGTTATCGTCGTCGTCACAAAGGAACCTGGAACCTTGAACCT 1452
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 116 pAsnProMetAlaThrGluIleGlyIleAlaIleGlnGlyGluGlyMetValArgVal 136
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1453 GGCTGTAAAGAAAGAGCAACACAGAGGGGACGCGGGAAGAGAGGAGCAAGACGA 1512
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 136 lCys----- 137
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1513 AGAAGAGGAGGAAGTTAACAGAGAGGTGCTAGTACACAGCGAGTTGAAGAGGCGA 1572
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 138 -SerLysSerGlyThrGlyCysLysAsnMetArgPhe-----LysValGluGluGly 155
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1573 TGTGTTTCATATGCCACGAGCTCATCCAGTAGGCATCAACGCT-----TCCTCCGACT 1626
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 155 pValPheValValProArgPheAspProMetAlaGlnMetAlaPheAsnAsnSerPh 175
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1627 CCATCTCTTGGGTTTCGTTCAACGCTGAAACCGTGAACACACAGAGATCTTCTTCG 1686
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
Db 175 eValPheValGlyPheSerThrThrThrLysLysHisHisProGlnTyrLeuThrGly-- 194
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1687 TAAGGACAAATGTATAGACCAGATAGAGAACGAAAGGATTTTAGCATTCCTCGGTC 1746
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 195 -LysAlaSerValLeuArgThrLeuAspArgGlnIleLeuGluAlaSerPheAsnValG 214
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1747 GGGTGAACAAGTTGAGAAGCTCATCAAAAAACCAAGGAATCTCACTTTGTGAGTGCTCG 1806
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 214 yAsnThrThrMetHisGlnIleLeuGluAlaGlnGlyAspSerValIleLeu----- 231
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1807 TCTCTCAATCTCAATCTCAATCTCGTCTCTCGAGAAAGAGTCTCTCGAGAAAGAGA 1866
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 232 -----GluCysThrSerCysAlaGluGluLysArgLeuMetGluGluGluMe 248
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 1867 TCAAGAGGAGGAAACCAAGAGGGAAG 1894
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 248 tArgLysGluGluGluAlaLysLys 257
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 10
US-10-784-004-709
; Sequence 709, Application US/10784004
; Publication No. US20060084066A1
; GENERAL INFORMATION:
; APPLICANT: Biogen Idec
; TITLE OF INVENTION: Surrogate Markers of Pain
; FILE REFERENCE: 08201.6029-00000
; CURRENT APPLICATION NUMBER: US/10/784,004
; CURRENT FILING DATE: 2004-02-20
; NUMBER OF SEQ ID NOS: 1251
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 709
; LENGTH: 5689
; TYPE: PRT
; ORGANISM: human
US-10-784-004-709

Alignment Scores:
Pred. No.: 0.000117 Length: 5689
Score: 162.00 Matches: 133
Percent Similarity: 32.0% Conservative: 87
Best Local Similarity: 20.4% Mismatches: 299
Query Match: 4.5% Indels: 211
DB: 9 Gaps: 28

US-10-728-323-1 (1-2032) x US-10-784-004-709 (1-5689)
QY 17 CATCAATCATCTATATAAGTAGTAGCAGGAGCAATGAGAGGGGTTTCTCACTGATG 76
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 2727 HisGlnGluSerAlaArgAspTrpSerGlyGlyArgSerGlyArgSerGlySerPhe 2746
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 77 CTGTTCTAGGGATCCTTCTCTCTGGCTTCAGTTTCTGCAACGATCCCAAGTCATCACCT 136
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 2747 TyrGlnVal-----SerThrHisGluGlnSerGluSer 2757
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 137 TACCAGAAGAAACAGAGAACCCCTCGGCCAGAGTGCTCCAGAGTTGTCAACAGGAA 196
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 2758 AlaHisGlyArgThrArgThrSerThrGlyArgArgGlnGlySerHisHisGluGlnAla 2777
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 197 CCGATGACTTGAAGCAAAAGGCATCGAGTCTCGTCGCCCAAGCTCGAGTATGATCCT 256
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 2778 ArgAspSerSerArgHisSerAlaSerGlnGluGlnAspThrIleArgAlaHisPro 2797
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 257 -----CGTTGTGCTCTATGATCCTCGAGGA 280
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 2798 GlySerArgArgGlyGlyArgGlnGlySerHisHisGlnSerValAspArgSerGly 2817
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 281 CACATCTGGC-----ACCAGCAACCAAC-- 301
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 2818 HisSerGlySerHisHisSerHisThrSerGlnGlyArgSerAspAlaSerHisGly 2837
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 302 -----CGTTCCCTCCAGGGGACGACACGTGGCCGCCCAACCCCGAGAC--- 346
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 2838 GlnSerGlySerArgSerAlaSerArgGlnThrArgLysAspLysGlnSerGlyAspGly 2857
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```



```
Percent Similarity: 34.3% Conservative: 83
Best Local Similarity: 18.8% Mismatches: 210
Query Match: 4.4% Indels: 140
DB: 9 Gaps: 19

US-10-728-323-1 (1-2032) x US-10-508-263-20 (1-495)

QY 467 AGTCATCAGACGACCGGAAA-----ATAAGGCC 496
||| :|||:|||||: :|||:
Db 22 SerArgGluGlnProGlnGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysPro 41
QY 497 GAAGGAGA---GAAGGAGAACAGAGTGGGGAACACCGAGGTAGCCATGTGAGGGAGAA 553
||| :|||:|||||: ||| :|||:
Db 42 AspAsnArgIleGluSerGlu-----GlyGlyLeuIleGluThrTrp 55
QY 554 ACATCTCGGAACAACACCTTTCTACTTCCCGTCAGGCGGTTTAGCACCCGCTACGGGAAC 613
||| :|||:|||||: ||| :|||:
Db 56 AsnProAsnAsnLysProPheGlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsn 75
QY 614 CAAACGGTAGGATCCGGGTCCTCGCAGAGTTTGACCAAGGTCAAGGCAGTTTCAGAAAT 673
||| :|||:|||||: ||| :|||:
Db 76 ArgAsnAla-----LeuArgArgProSerTyrThrAsn 86
QY 674 CTCAGAAATCACCGTATTGTGCAGATCGAGGCCAAA----- 709
||| :|||:|||||: ||| :|||:
Db 87 GlyProGlnGluIleTyrIleGln---GlnGlyLysGlyIlePheGlyMetIleTyrPro 105
QY 710 -----CCTAACACTCTGTCTTCCCAAGCAGCGTGTGCTGATAACATCTCTGTATC 763
||| :|||:|||||: ||| :|||:
Db 106 GlyCysProSerThrPheGluGluProGlnGlnProGln----- 118
QY 764 CACAGAGGCAACCCACCGTAGCGTAGCAATAGGCAATAACAGAAAGACTTTAATCTT 823
||| :|||:|||||: ||| :|||:
Db 119 GlnArgGlyGlnSer-----SerArgProGlnAspArgHisGlnLysIleTyrAsnPhe 136
QY 824 GACGAGGGCCATGCACATCCAGATCCCATCCGTTTCATTTCTTACATCTTGAACCCCAT 883
||| :|||:|||||: ||| :|||:
Db 137 ArgGluGlyAspLeuIleAlaValProThrGlyValAlaTrpIleMetTyrAsnAsnGlu 156
QY 884 GACAACGAGAACCTCAGAGTAGCTAAATC----- 913
||| :|||:|||||: ||| :|||:
Db 157 AspThrProValValAlaValSerIleIleAspThrAsnSerLeuGluAsnGlnLeuAsp 176
QY 914 TCATGCGCGTTTAACACACCCGCGCAGTTTGAGGATTTCTCCCGGAGCAGCCAGAC 973
||| :|||:|||||: ||| :|||:
Db 177 GlnMetPro-----ArgArgPheTyrLeuAlaGlyAsnGlnGlu 189
QY 974 CAA----- 976
||| :|||:|||||: ||| :|||:
Db 190 GlnGluPheLeuLysTyrGlnGlnGlnGlnGlyHisGlnSerGlnLysGlyLysHis 209
QY 977 -----TCATCTACTTCGAGGCTTCAGCAGGATACGTTG 1012
||| :|||:|||||: ||| :|||:
Db 210 GlnGlnGlnGluGluAsnGluGlyGlySerIleLeuSerGlyPheThrLeuGluPheLeu 229
QY 1013 GAGCGCGCTTCAATCGCGAATTCAATGACATAGCGGGTGTCTTAGAGAGATGCA 1072
||| :|||:|||||: ||| :|||:
Db 230 GluHisAlaPheSerValAsp---LysGlnIleAlaLysAsnLeuGlnGlyGluAsnGlu 248
QY 1073 GGAGGTGAGCAA----- 1084
||| :|||:|||||: ||| :|||:
Db 249 GlyLysAspGlyAlaIleValThrValLysGlyLeuSerValIleLysProPro 268
QY 1085 ---GAGGAGAGAGGCGAGCGGATGGAGTACTCGGAGTAGTGAGAACAAATGAAGAGTG 1141
||| :|||:|||||: ||| :|||:
Db 269 ThrAspGluGlnGlnGlnArgProGlnGlnGluGluGluGluGluGluGluLysPro 288
QY 1142 ATAGTCAAGTGTCAAGGAGCAGCTTGAGAACTTACTAAGCAGCGTAATCGCTCA 1201
||| :|||:|||||: ||| :|||:
Db 289 GlnCysLysGlyLysAspLysHisCysGlnArg---ProArgGlySerGlnSerLysSer 307
QY 1202 AAGAAGGCTCCGAAGAAGAGGAGGATATCACCAACCAATCACTTGAGAGAGCGGAG 1261
||| :|||:|||||: ||| :|||:

Db 308 ArgArgAsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIleGlyGln 327
QY 1262 -----CCGATCTTTCTAACAACTTTGGGAAGTTATTGAGGTGAAGCCAGCAAG 1312
||| :|||:|||||: ||| :|||:
Db 328 ThrSerSerProAspIleTyrAsnProGlnAlaGlySerValThrThrAlaThrSerLeu 347
QY 1313 AAGAACCACCGCTTCAGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAAGGA 1372
||| :|||:|||||: ||| :|||:
Db 348 AspPheProAlaLeuSerTrpLeuArgLeuSerAlaGluPheGlySerLeuArgLysAsn 367
QY 1373 GCTTTGATCTCCACACTTCAACTCAAAAGGCCATGGTTATCTGCTCGCTCAACAAGGA 1432
||| :|||:|||||: ||| :|||:
Db 368 AlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeuAsnGly 387
QY 1433 ACTGGAAACCTTGAACCTCGTGTGTAGAAAAGACAAACAAGAGGGGACGCGGGAA 1492
||| :|||:|||||: ||| :|||:
Db 388 ArgAlaLeuIleGlnValValAsnCysAsnGlyGlu----- 399
QY 1493 GAAGAGGAGACGACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1552
||| :|||:|||||: ||| :|||:
Db 400 -----ArgValPheAsp 403
QY 1553 GCGAGGTTGAAGAAAGCGCATGTGTTTCATGCTCCAGCAGCTCATCCAGTAGCCATCAAC 1612
||| :|||:|||||: ||| :|||:
Db 404 GlyGluLeuGlnGluGlyArgValLeuIleValProGlnAsnPheValValAlaAlaArg 423
QY 1613 GCTTCTCCGAA---CTCCATCTGCTTGGCTTCGGTATCAACGCTGAAAACACACAGAGA 1669
||| :|||:|||||: ||| :|||:
Db 424 SerGlnSerAspAsnPheGluTyrValSerPheLysThrAsn-----AspThrProMet 441
QY 1670 ATCTTCTTCCAGGTGATGAAGGACATGTGATAGCCACATAGAGAGGAGGAGGAGGAGGAT 1729
||| :|||:|||||: ||| :|||:
Db 442 IleGlyThrLeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluValIleGln 461
QY 1730 TTAGCATTCCTCCGCTCGGCTGAAACAAGTTGAGAGCTCATCAAAAACCAAGGAATCT 1789
||| :|||:|||||: ||| :|||:
Db 462 HisThrPheAsnLeuLysSerGlnGlnAlaArgGln---IleLysAsnAsnAsnProPhe 480
QY 1790 CACTTTGTGCTGCTCGTCTCTCAATCTCAATCTCAATCT 1828
||| :|||:|||||: ||| :|||:
Db 481 LysPheLeuValProProGlnGluSerGlnLysArgAla 493

RESULT 12
US-10-982-545-12
; Sequence 12, Application US/10982545
; Publication No. US20050244890A1
; GENERAL INFORMATION:
; APPLICANT: Davies, Huw Alun
; APPLICANT: McGuire, James
; APPLICANT: Simonsen, Anja Hviid
; APPLICANT: Blennow, Kaj
; APPLICANT: Podust, Vladimir
; APPLICANT: CIPHERGEN Biosystems, Inc.
; TITLE OF INVENTION: Biomarkers for Alzheimer's Disease
; FILE REFERENCE: 016866-011550US
; CURRENT APPLICATION NUMBER: US/10/982,545
; PRIOR FILING DATE: 2004-11-06
; PRIOR APPLICATION NUMBER: US 60/518,360
; PRIOR FILING DATE: 2003-11-07
; PRIOR APPLICATION NUMBER: US 60/526,753
; PRIOR FILING DATE: 2003-12-02
; PRIOR APPLICATION NUMBER: US 60/546,423
; PRIOR FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: US 60/547,250
; PRIOR FILING DATE: 2004-02-23
; PRIOR APPLICATION NUMBER: US 60/558,896
; PRIOR FILING DATE: 2004-04-02
; PRIOR APPLICATION NUMBER: US 60/572,617
; PRIOR FILING DATE: 2004-05-18
; PRIOR APPLICATION NUMBER: US 60/586,503
; PRIOR FILING DATE: 2004-07-08
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
```

```

; SEQ ID NO 12
; LENGTH: 677
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Chromogranin B precursor
; NAME/KEY: SIGNAL
; LOCATION: (1)..(20)
; OTHER INFORMATION: signal peptide
; NAME/KEY: PEPTIDE
; LOCATION: (21)..(677)
; OTHER INFORMATION: Chromogranin B (Secretogranin I)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (326)..(385)
; OTHER INFORMATION: biomarker peptide 7258 Da, processed fragment of
; OTHER INFORMATION: Chromogranin B
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (341)
; OTHER INFORMATION: Xaa = sulfotyrosine
US-10-982-545-12

Alignment Scores:
Pred. No.: 0.000274 Length: 677
Score: 153.00 Matches: 143
Percent Similarity: 34.3% Conservatives: 101
Best Local Similarity: 20.1% Mismatches: 285
Query Match: 4.3% Indels: 183
DB: 9 Gaps: 34

US-10-728-323-1 (1-2032) x US-10-982-545-12 (1-677)
QY 71 CTGATGCTGTTCTAGGATCTTCTGCTGGCTTCAGTTCTGCA----- 115
Db 6 LeuLeuSerLeuLeuGlyAlaValGlyLeuAlaValAsnSerMetProValAspAsn 25
QY 116 -----ACGATGCTCAAGTCATCACCTTAC 139
Db 26 ArgAsnHisAsnGluGlyMetValThrArgCysIleIleGluValLeuSerAsnAlaLeu 45
QY 140 CAGAGAAGAAACAGAGAACCCCTGCGCCAGAGTGCTCCAGAGTTGTCAACAGGAACCG 199
Db 46 SerLysSerSerAlaProProlleThrProGluCysArgGlnValLeuLysThrSerArg 65
QY 200 GATGACTTGAACAAAG---GCATGCGAGTCTCGCTGCACCAAGCTCGATGATCATCT 256
Db 66 LysAspValLysAspLysGluThrThrGluAsnGluAsnThrLysPheGluVal----- 83
QY 257 CTTTGTGTATGATCTCTCGAGGACACACTGGCCACCACCAACCAACGTTCCCTCCAGG 316
Db 84 ArgLeuLeuArgAspPro---AlaAspAlaSerGluAlaHisGluSerSerSerArgGly 102
QY 317 GAGCGGACACGTGGCGCCCAACCCGAGACTACGATGATGACCGCGCTCAACCCCGAAGA 376
Db 103 Glu-----AlaGlyAlaProGly-----GluGluAspIleGlnGlyProThrLys 117
QY 377 GAGGAAGGCGCGATGGGACACAGCTGGACCGGAGGCGGTGAAGA-----GAAGAA 430
Db 118 AlaAspThrGluLysTrpAlaGluGlyGlyHisSerArgGluArgAlaAspGluPro 137
QY 431 GACTCG-----AGACAACCAAGAGAGATTGGAGGCGACCAAGTCAT 472
Db 138 GlnTrpSerLeuTyProSerAspSerGlnValSerGluGluValLysThrArgHisSer 157
QY 473 CAGCAGCCACCGAAATAGGCCCGCAAGGAAGAGAGAGAGAAACA-----GAG 520
Db 158 GluLysSerGlnArgGluAspGluGluGluGluGluGlyGluAsnTyGlnLysGlyGlu 177
QY 521 TGGGGAACACCGAGGTAGC-----CATGTAGGGGAAGAAACATCTCGGAACACCT 571

```

```

Db 178 ArgGlyGluAspSerSerGluGluLysHisLeuGluGluProGlyGluThrGlnAsnAla 197
QY 572 TTCTACTTCCCGTCAAGCGCGTTTAGCACC----- 601
Db 198 PheLeuAsnGluArgLysGlnAlaSerAlaIleLysLysGluGluLeuValaAlaArgSer 217
QY 602 -----CGCTACGGAAACAAACCGTAGGATCCGGGTCTCTGCAGAGGTTTGACCAA 652
Db 218 GluThrHisAlaAlaGlyHisSerGlnGluLysThrHisSerArgGluLysSerGln 237
QY 653 AGGTCAAGCGCAGTTTCAGAATCTCCAGAATCACCGTATTGTGCACATCGAGGCCAACCT 712
Db 238 GluSerGlyGluGluAlaGlySerGlnGluAsnHisProGlnGluSerLysGlyGlnPro 257
QY 713 AACACTCTTGTCTTCCCAAGCAGCTGATGCTGATAACATCTCTTGTATTATCCAGCAAGG 772
Db 258 ArgSerGlnGluGluSerGluGluGlyGluGluAlaThrSerGluValAspLysArg 277
QY 773 CAAGCCACCGTACCGTAGCAATGGCAATTAACAGAAAGAGCTTTAACTTTGACGAGGCG 832
Db 278 ArgThrArgProArgHisHisGlyArgSerArgProAspArgSerSerGlnGlyGly 297
QY 833 CATGCACCTCAGATCCCATCCGTTTCATTCTTCATCTTCAACCGCCATCACAACCCAG 892
Db 298 -----SerLeuProSerGluGluLysGlyHisProGlnGluGluSerGluSer 314
QY 893 AACCTCAGATAGTCTAAATCTCCATGCCCGTTAAACACACCCGCGCAGTTTGAAGATTTC 952
Db 315 AsnValSerMetAlaSerLeu-----GlyGluLysArgAspHis 327
QY 953 TTCCCGCGCAGCAGCCGA-----GACCAATCATCTTACTTTCGAGGCG 994
Db 328 HisSerThrHisTyArgAlaSerGluGluGluProGlu**GlyGluGluLysGly 347
QY 995 TTCAGCAGAATACGTTGAGCGCCCTTCAATCGGAATCAATGAGATCAGACGAGGTTG 1054
Db 348 TyrPro-----GlyValGlnAlaProGluAspLeuGluTrpGluArgTyrArg----- 363
QY 1055 CTGTTAGAAGAGAATGCGAGGAGTGAGCAAGAGAGAGAGAGGCGCAGG----- 1102
Db 364 -----GlyArgGlySerGluGluTyArgAlaProArgProGlnSerGlu 378
QY 1103 ---CGATGAGTACTCGGAGTAGTGAGAAACAATGAGGAGGATAGTCAAGTGTCAAAG 1159
Db 379 GluSerTrpAspGluGluAspLysArgAsnTyProSerLeu----- 392
QY 1160 GAGCAGCTTGAAGAACTTACTAAGCACCGCTAAATCCGCTCAAGAAAGGCTCCGAAGA 1219
Db 393 -----GluLeuAspLysMetAlaHisGlyTyGlyGluGlu---SerGluGlu 407
QY 1220 GAGGAGATATCACCAACCAATCAACTTGAGAGAAGCGCAGCCGATCTTTCTAACAC 1279
Db 408 Glu--ArgGlyLeuGluProGlyLysGlyArgHisHisArgGlyArgGly----- 423
QY 1280 TTTGGGAAGTTATTGAGTGAAGCCAGA-----CAAGAAGAAC 1318
Db 424 -----GlyGluProArgAlaTyProPheMetSerAspThrArgGluGlu 438
QY 1319 CCCAGCTTTCAGGACCTGACATGATGCTCACCTGTGTAGAGATCAAGAGAGGAGCTTTG 1378
Db 438 ysaRgPheLeuGluGlyGlyHisHis-----ArgValGlnGluAsnGlnMetA 454
QY 1379 ATGCTCCACACTTCAACTCAAGGCCATGGTTATCTGCTCGTCGTCACCAAGAGAACTGA 1438
Db 454 spLysAla-----ArgArgHisProGlnGlyAlaTrpL 465
QY 1439 AACCTTGAACCTCGT-----GGCTGTGAAGAAAGAGCAACAA 1474
Db 465 ysaGluLeuAspArgAsnTyLeuAsnTyGlyGluGluGlyAlaProGlyLys--TrpGln 484
QY 1475 CAGAGGCGACCGCGGGAAGAGAGAGGACCAAGCAAGAGAGAGAGAGGAGGAGTAAACAGA 1534
Db 485 GlnGlnGlyAspLeuGlnAspThrLysGluAsnArgGluGluAlaArgPheGlnAspLys 504

```

QY 1535 GAGGTG---CGTAGGTACACAGG-----AGTTGAAGAGGCGCATGTGTTTCATCATG 1585
Db 505 GlnTyrSerSerHisHisThrAlaGluLysArgLysArgLysGluGluLeuPhe---Asn 523
QY 1586 CACGACGCTCATCCAGTAGCCATCAACGCTTCCTCCGAACTCCATCTGCTTGGCTTCGGT 1645
Db 524 ProTyrTyrAspProLeuGlnTrpLysSerSer-----HisPhe----- 536
QY 1646 ATCAACGCTGAAACCAACCAAGAATCTTCCTTGCAGGTGATGAAGCAACATGTGATAGAC 1705
Db 537 GluArgArgAspAsnMetAsnAspAsnPheLeuGluGlyGluGluGluAsnGluLeuThr 556
QY 1706 CAGATAGACAAG----- 1717
Db 557 LeuAsnGluLysAsnPhePheProGluTyrAsnTyrAspTrpTrpGluLysLysProPhe 576
QY 1718 -----CAAGCGAAGGATTAGCAATTCCTGGTCCGGTGAA 1753
Db 577 SerGluAspValAsnTrpGlyTyrGluLysArgAsnLeuAla----- 590
QY 1754 CAAAGTTGAGAAGTCT---ATCAAAAAACCAAGAGAA-----TCTCACTTT 1795
Db 591 ArgValProLysLeuAspLeuLysArgGlnTyrAspArgValAlaGlnLeuAspGlnLeu 610
QY 1796 GTGAGTGTGCTCTCAATCTCAATCTCAATCTCCGTCGTCTCTGAGAAAGAGTCTCT 1855
Db 611 LeuHisTyrArgLysLysSer---AlaGluPheProAspPheTyrAspSerGluGluPro 629
QY 1856 GAGAAAGAGGATCAAGAGGAGGAGAAACCAA 1885
Db 630 ValSerThrHisGlnGluAlaGluAsnGlu 639

RESULT 13

US-11-072-512-3378
; Sequence 3378, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR FILING DATE: 2002-01-25
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3378
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-072-512-3378

Alignment Scores:

Pred. No.: 0.000566 Length: 623
Score: 149.00 Matches: 134
Percent Similarity: 34.2% Conservatives: 97
Best Local Similarity: 19.9% Mismatches: 274
Query Match: 4.2% Indels: 170
DB: 11 Gaps: 34
US-10-728-323-1 (1-2032) x US-11-072-512-3378 (1-623)
QY 116 ACGCATGCCAAGTCATCATCCCTTACCAGAAAGAAAACAGAACCCCTGCCCCAGAGGTGC 175
Db 5 SerHisPheSerGlnProProTyrArgAspLeuTrpProProArgProGlyGlyGlu 24
QY 176 CTCACAGAGTTGTCACAG-----GAACCGGATGACTTGAAGCAAAAGGCATCC--- 223
Db 25 ArgGluSerThrGlnArgLeuGlyGlnArgSerGlyAlaAspSerThrAlaCysSer 44
QY 224 -----GAGTCTCGCTCGCCACCAAGCTCGAGTATGATCCTCTGTT 262
Db 45 ArgAlaGlyThrProGlyAlaGluSerGluAlaGlyAlaCysTrpLeuHisProHisCys 64
QY 263 GTCTATGATCTCT-----CGAGGACACACTGGC 289
Db 65 SerPheThrProArgProArgArgGlyCysSerAspSerLeuArgGlySerArgSer 84
QY 290 ACCACCAACCAAGTTCCTCCAGGGAGCGGACACAGTGGCGCCCAACCCGGAGACTAC 349
Db 85 LeuSerAspValAlaArgArgProLeuGluArgSerArgLysHisArgPro----- 101
QY 350 GATGATGACCGCGTCAACCCCGAAGAGAGAGGAGGCGCGATGGGACAGCTGGACCG 409
Db 102 -----ArgSerArgArgLeuGluAspAla-----TrpGlyGluThrGlyThr 115
QY 410 AGGAGCGGTGAAAGAGAGAGACTGGAGACAAACAGAGAGAGATGGAGGCGACCAAGT 469
Db 116 Lys-----ProArgProAlaTrpGlnProGlnThr 125
QY 470 CATCAGCAGCCACGGAAATATAGGCC----- 496
Db 126 GlnLeuProGlnArgProGlnProCysProHisTyrProLeuAlaGlnGlyAspSer 145
QY 497 -----GAAGGA 502
Db 146 ProProProCysProGlyGlyAlaGlyThrProLeuSerGlyThrPheArgValGluLys 165
QY 503 AGAAGAGAGAACAGAGTGGGGAACACCA---GGTAGCCATCTGAGGAGAAACATCT 559
Db 166 AlaGlnGlyGlyAspGlnTrpAlaValProLeuGlyArgHisLeu-----Gly 181
QY 560 CGGAACAACCTTCTTCTTCTCCGTCAGCGGCTTTAGCACC-----CGTACGGG 610
Db 182 ArgTrpSerProSerSerValProSerGluArgSerSerValProSerGlnLysPheLys 201
QY 611 AACCAAAACGGTAGGATCCGGGTCCTGCAGAG-----TTTGACCAA----- 652
Db 202 ArgHisSerAlaCysValCysAlaGlnLysArgAspSerSerAspGlnValGluSerLeu 221
QY 653 AGGTCAAGGAGTTTCAGAAATCTCCAGAAATCCACCGTATGTGTCAGATCGAGGCCAACCT 712
Db 222 AlaSerArgAspSerGlnProLeuAlaSerSerLysGluMetArg-----SerPro 238
QY 713 AACACTCTTCTCTCCCAAGCAGCGTGTGATGATAACATCTCTTATATCCAGCAGGG 772
Db 239 HisThrGlnValLeu-----LysSerLysLeuGluGluValValSerSerGlnAsp 256
QY 773 CAAGCCACCGTCGACCGTAGCAATGGCAATAACAGAAAGAGCTTTAATCTTGTACGAGGC 832
Db 257 GlnGlnIleValAlaLeuValLeuThrArgLeuLysLysAlaGlnArgIleArgGlu--- 275
QY 833 CATGCACTAGAAATCCATCCCGGTTTCATTTCTCATCTTTGACCGCCATGACACACAG 892
Db 276 -----LeuGlnGlnAlaAlaLysAlaTrpGluGluLysArgSerAspGln--- 292


```
QY 893 AACCTCAGAGTAGCTAAATCTCCATGCCCCGTTAAACACACCCCGCCAGTTTGAGGATTTTC 952
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
293 -----LysValGlnMetThrLeuGluArgLeu----- 304
QY 953 TTCCCGCGCAGCAGACCAATCATCTCTACTTCAGCGGCTTCAGCAGGATACGTTG 1012
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
305 -----LeuLeuArgGlnSerGlnGluIntrpGlnGluLysGlnArgLysThrLeu 322
QY 1013 GAGGCGCGCTTCAATCGGAATTCATAGATACGAGGGTGTGTGTAGAAAGAGATGCA 1072
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
323 Gln-----SerProGluGlnArgGlyLeuArgAspSerGlnArgLysAsnVal 339
QY 1073 GGAGGTGACAAAGAGAGAGAGCGCAGAGCGATGAGTACTCGGAGTAGTGAGAACAAAT 1132
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
340 -----ProProGlyGluSerArgTTPLeuGluGlnProGluAspGlnGlu 354
QY 1133 GAAGGAGTGATAGTCAAGTGTCAAGGAGCAGCGTTGAAGAATCTACTAAGCAGCGTAA 1192
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
355 SerProArgGlnGluLysLeuGluLysAlaAaGAlaGln-----AlaGluHisArgLys 372
QY 1193 TCCGTCTCAAAAGAGGCTCCGAAGAAGAGGAGATATCACCAACCAATCAACTTGAGA 1252
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
373 GlnCysGlnValArgLeuArgGluGlnGluLysMetLeuArg-----AsnLeuArg 390
QY 1253 GAAGCGAGCCCC--GATCTTTCTAAACAACTTTTGGGAAG-----TTA 1291
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
391 GluGlnHisSerLeuGlnLeuArgArgLeuValGluAlaCysArgLysArgHisLeu 410
QY 1292 TTTGAGGTGAAGCCACAGAAGAAGACCCAGCTTCAGGACCTGGACATG----- 1342
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
411 HisAlaValGluGlyGlnLysLys-----ValGlnAspThrAsnLeuSerSerLeu 427
QY 1343 -----ATGCTCACCTGTGTAGATCAAAAGAGGAGCTTTG 1378
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
428 IleAsnTyrGlnAlaArgLysValLeuMetAspCysGlnAlaLysAlaGluLeuLeu 447
QY 1379 ATGCTCCACACTTCAACTCAAGGCCATCGTTATCGTCTGCTCAACAAAGAACTGGA 1438
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
448 ArgGlnLeuSerLeuGluGlnSerPheGlnArgSerGlnGluLeuLeu----- 465
QY 1439 AACCTTGAAGCTGTGCTTAAGAAAGACCAACACAGAGGGGCGGCGGAAGAGAG 1498
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
466 -----LeuArgLysGluArgGlnArgLeuArgGluLeuArgGluLysAlaGln 479
QY 1499 GAGGACGAAGACGAAGAGAGAG-----GGAAGTAACAGAGGTGCGT 1543
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
480 LysGluGluGluGlnLeuGlnAlaArgTTPArgAlaGlyGluSerGluGlnArg 499
QY 1544 AGGTACACAGCGAGTTG-----AAGGAAGCGATGTGTTTCATATGCCAGCA--GCT 1594
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
500 LysMetArgLysArgGluLeuValGluLeuAlaAspGluLysIleArgGlnAlaArgSer 519
QY 1595 CATCCAGTAGCCATCAACGCTTCTCCGAACTCCATCTGCTTGGCTTGGGTATCAACGCT 1654
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
520 HisValHisLysThrThrArgAspLysValGlnHisLeuArgGluLeuAsnHisLeuArg 539
QY 1655 GAAACCAACCAACAGAGATTTAGCATTTCCCTGGTGGGTGAACAAAGTTGAGAGCTCATCAA 1714
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
540 GluLysAsnHisIleLeuLysLeu-----LysAlaGlu 551
QY 1715 AAGCAAGCAAGAGATTTAGCATTTCCCTGGTGGGTGAACAAAGTTGAGAGCTCATCAA 1774
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
552 LysGluGluLys---CysHisIleGluGlyIleLysGluAlaIleLysLys-----Lys 568
QY 1775 ACCAGAAGCAATCTACATTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCGGTG 1834
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
569 GluGlnArgValGlnHisIle-----SerGlnGlyLysAspProAsn 582
QY 1835 TCTCTCTAGC---AAAGAGTCTCTGAGAAAGAGGATCAAGAGAG 1876
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
583 PheGlnGluPheGlnLysLeuProGlnAlaSerArgArgGluGlu 597
```

RESULT 14

```
US-10-508-263-24
; Sequence 24, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Glycine max
US-10-508-263-24
```

```
Alignment Scores:
Pred. No.: 0.000656 Length: 562
Score: 148.00 Matches: 68
Percent Similarity: 38.8% Conservative: 45
Best Local Similarity: 23.4% Mismatches: 129
Query Match: 4.1% Indels: 49
DB: 9 Gaps: 11
```

US-10-728-323-1 (1-2032) x US-10-508-263-24 (1-562)

```
QY 292 CACCAACCAACGTTCCCTCCAGGGGAGCGGACACGTGGCCGCCCAACCCGGAGACTACCA 351
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
298 HisProPProArgArgProSerHisGly-LysArgGluGlnAspGluAspGluAspGluAs 317
QY 352 TGATCACCCTGCTCAACCCCGAAGAGGAGGAGCGCGATGGGACCACAGCTGGACCGAG 411
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
317 pGluAspLysProArgProSerArgProSerGlnGlyLysArgAsnLysThrGlyGlnAs 337
QY 412 GGAGCGTGAAGAGAAAGAA-----GACTGGAGACAACCAAG 447
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
337 pGluAspGluAspGluAspGluAspGlnProArgLysSerArgGluTrpArgSerLysLy 357
QY 448 AGAAGATTGGAGGCGCACCAAGTCATCAGCAGCCACGGAAATAAGGCCCGAAGGAGACA 507
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
357 sThrGlnProArgArgProArgGlnGluProArgGluArgGlyCysGluThrArgAs 377
QY 508 AGGAAACAAGAGTGGGAAACACACAGGTAGCCATGTAGGGGAAGAAACATCTCGGAACAA 567
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
377 nGlyValGluGlu---AsnIleCysThrLeuLysLeuHisGluAsnIleAlaArg----- 394
QY 568 CCCTTTTACTTCCGTCGAAGCGGTTTAGCACCCGCTACGGGAACCAAAACGGTAGGAT 627
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
395 -----ProSerArg-----AlaAspPheTyrAsnProLysAlaGlyArgI 408
QY 628 CCGGTCTCTCAGAGGTTTGACCAAGGTCA---AGGCAGTTTTCAGATCTCCAGAATCA 684
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
408 sSerThrLeuAsnSerLeuThrLeuProAlaLeuArgGlnPheGlnLeuSerAlaGlnTy 428
QY 685 CCGTATTGTGAGATCGAGGCGCAACACTTGTCTTCTTCCCAAG---CACGCTGA 741
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
428 rValValLeuTyr-----LysAsnGlyIleTyrSerProHisTrpAsnLeuAs 444
QY 742 TGCTGATTAACATCTTGTATTATCCAGCAAGGGCAAGCCCGCTAGCCGTAGCAAAAT----- 796
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
444 nAlaAsnSerValIleTyrValThrArgGlyGlnGlyLysValArgValValAsnCysG 464
QY 797 -GGCAATAACAGAGAGCTTTAATCTTTCAGCGAGGCGCATGCACATCAGATCCCATCCCG 855
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
464 nGlyAsnAlaValPheAspGlyGluLeuArgArgGlyGlnLeuLeuValValProGlnAs 484
QY 856 TTTTCATTCTTACATCTTGAACCGCCATGACAAACCAACCACTCAGAGTAGCTAGCTAAAATCTC 915
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
484 nPheVal-----ValAl 488
QY 916 CATGCCCGTTAACACACCGCGCAGTTTGTAGGATTTCTTCCGCGGAGAGCGCGAGACCA 975
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
```

Db 488 aGluGlnAlaGlyGluGlnGlyPheGluTyrIleValPheLysThrHisAsnAlaVa 508
 QY 976 ATCATCTCTATTGCAGCGCTTCAGCAGG-----AATACCTGTGGAGCGCCGCTT 1023
 Db 508 lThrSerTyrLeuLysAspValPheArgAlaIleProSerGluValLeuAlaHisSerTy 528
 QY 1024 CAATCGCGAATTCAATAGATACGAGAGGTGCTGTGTAGAGAGAATGCAGGA----- 1075
 Db 528 rAsnLeuArgGlnSerGlnValSerGluLeuLysTyrGluGlyAsnTrpGlyProLeuVa 548
 QY 1076 -GCTGAGCAAGAGGAGAGGCGCAGAGGCGCA 1105
 Db 548 lAsnProGluSerGlnGlnGlySerProArg 558

RESULT 15

US-11-072-512-3385
 ; Sequence 3385, Application US/11072512
 ; Publication No. US20060029945A1

GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO
 ; APPLICANT: SUGIYAMA, TOMOYASU
 ; APPLICANT: OTSUKI, TETSUJI
 ; APPLICANT: WAKAMATSU, AI
 ; APPLICANT: SATO, HIROYUKI
 ; APPLICANT: ISHII, SHIZUKO
 ; APPLICANT: YAMAMOTO, JUN-ICHI
 ; APPLICANT: ISONO, YUUKO
 ; APPLICANT: HIO, YURI
 ; APPLICANT: OTSUKA, KAORU
 ; APPLICANT: NAGAI, KEIICHI
 ; APPLICANT: IRIE, RYOTARO
 ; APPLICANT: TAMECHIKA, ICHIRO
 ; APPLICANT: SEKI, NAOHICO
 ; APPLICANT: YOSHIKAWA, TSUTOMU
 ; APPLICANT: OTSUKA, MOTOYUKI
 ; APPLICANT: NAGAHARI, KENJI
 ; APPLICANT: MASUHO, YASUHIKO
 ; TITLE OF INVENTION: Novel full length cDNA
 ; FILE REFERENCE: 084335-0191
 ; CURRENT APPLICATION NUMBER: US/11/072,512
 ; CURRENT FILING DATE: 2005-03-07
 ; PRIOR APPLICATION NUMBER: US 60/350,978
 ; PRIOR FILING DATE: 2002-01-25
 ; PRIOR APPLICATION NUMBER: JP 2001-379298
 ; PRIOR FILING DATE: 2001-11-05
 ; NUMBER OF SEQ ID NOS: 4096
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3385
 ; LENGTH: 715
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens

US-11-072-512-3385

Alignment Scores:
 Pred. No.: 0.000794 Length: 715
 Score: 147.50 Matches: 141
 Percent Similarity: 34.9% Conservative: 101
 Best Local Similarity: 20.3% Mismatches: 275
 Query Match: 4.1% Indels: 176
 DB: 11 Gaps: 30

US-10-728-323-1 (1-2032) x US-11-072-512-3385 (1-715)

QY 95 GTCTGTGCTTCAGTTCTGCAACGCGATGCCAAGTCATCACCTTACCAGAGAAGAAACAGAG 154
 Db 56 ValMetSerProSerSerLysSerThrLysSerThrLysSerThrLysArgAlaProSer 75
 QY 155 AAC---CCCTGCGCCAGAGGTGCCTCCAGAGTTGTCAACAGGACCGGATGACTTGAAG 211
 Db 76 AsnArgProSerSerArgValArgSerLysAlaArgValArgThrProSerArgValSer 95
 QY 212 CAAAGGCATCGAGTCTCGCTGCACCAAGCTCGAGTATGATCTCTCGTGTGTCTAT--- 268

Db 96 ThrAspThrArgThrSerLysAlaSerLysAlaSer---AspValArgCysHisGlnArg 114
 QY 269 -----GATCCTCGAGGACACACT-----GGCACCCACCAACCAACGCT 304
 Db 115 ArgGlyThrHisSerArgGlyArgThrProGlyArgArgGlySerArgSerSerLysArg 134
 QY 305 TCCCTC-----CCAGGGAG---CGGACACGTGGCGCCCAACCCGAGAC 346
 Db 135 SerProSerArgAlaSerThrProGlyArgIleArgThrHisGlyAlaArgProGlyMet 154
 QY 347 TAGCATGATACCGCGCTCAACCCCGAAGAGAGGAGGCGGATGGGAGCAGCTGGA 406
 Db 155 AlaSerArgValArgThrProThrSerGlnGlnLysGlySerArgGlyLysSerTyrGly 174
 QY 407 ---CCGAGGAGCGTGAAGAGAAGACTCGAGACAACCAAGAGAAGATTGGAGCGCA 463
 Db 175 ArgProArgThrSerAsnArgGluArgSerAspSerGlnProArg----- 189
 QY 464 CCAAGTCATCAGCAGCCCGGAAATAAGGCCCGAAGGAGGAGAGAGAGGA----- 511
 Db 190 ---AsnLeuSerLysSerTyrArgProGlyGlySerGlyIleGlyArgSer 207
 QY 512 -----GAACAAGAGTGGGGAACACCA---GGTAGCCATGTG 544
 Db 208 SerGluLeuAlaValThrProSerThrAlaLysCysGlnThrProThrGlyIleProSer 227
 QY 545 AGGGAAGAAACATCTCGGAACACCT----- 571
 Db 228 LysGluLysSerAspAsnProSerProSerSerArgLysValLysSerTyrGlyGln 247
 QY 572 TTCTACTTCCCGTCAGGCGGTTTAGCACCCGCTACCGGAACCAACCGTAGTCGCG 631
 Db 248 MetIleIleProSerArgGluLysSerTyrSerProThrGluMetSerArgValLys 267
 QY 632 GTCCTGCAGAGGTTTGACCAAGGTCAAGGCAGTTTTCAGAAATCTCCAGAAATCACCGTATT 691
 Db 268 SerTyrAsnGlnAlaSerThrArgSerArgPro-----GlnSerHis----- 281
 QY 692 GTGCAGATCAGGCCAAACCTAACACTCTTGTTCTTCCCAAGCACCGCTGATGTAAC 751
 Db 281 ----- 281
 QY 752 ATCCTTGTTATCCAGCAAGGCAAGCCACCGTCACCGTAGCAATGCAATAACAGAAAG 811
 Db 282 -----SerGlnSerArgSerProArgSerArgSerGlySerGlnLysArg 297
 QY 812 AGCTTTAATCTTGACGAGGCGCATGCA-----CTCAGA 844
 Db 298 ThrHisSerArgValArgSerHisSerTyrLysArgAsnHisSerArgAlaArgSerArg 317
 QY 845 ATCCCATCCGTTTCATCTTACATCTTGAACCGCCATGACACCAACGACACCTCAGAGTA 904
 Db 318 ThrArgLysGlyIleLeuSerGlnMetGlyArgHisSerGlnSerArgSerHisSerLys 337
 QY 905 GCTAAATCTCCATGCCCGTTAAACACACCC-----GGCAGTTTGAGGATTTCTCCCG 958
 Db 338 GlyLysSerGlnAsnGlnSerArgThrProArgArgGlyArgSerHisAsnTrpSerArg 357
 QY 959 GCGAGCAGCGAGACCAATCATCTACTTGCAGGGGTTTCAGCAGGAATACGTTGGAGGCC 1018
 Db 358 AsnProSerLysGluArgSer-----HisSerHisSerArgSerSerLysGlu 374
 QY 1019 GCCTTCAATCGGAATTCAATGAGATACGAGGGTCTGTTAGAAGAGATGACGAGGT 1078
 Db 375 ArgAspHisArgGlySerSerProArgLysGluSerGlyArgSerGlnSerGlySer 394
 QY 1079 GACCAAGAGAGAGCGGAGCGATGGAGTACTCGGAGTAGTAGAGAAACAATGAAGGA 1138
 Db 395 ProAsnLysGlnArgAspHisSerArgSerArgSerProAsnLysAlaArgAsp--ArgS 414
 QY 1139 GTGATAGTCAAAGTGTCAAAGGAGCACGTTGAAGAACT----- 1176
 Db 414 erArgSerArgSerProTyrLysAlaArgAspArgSerArgSerArgSerProAsnLysA 434

```
QY 1177 -----TACTAAGCAGCTAAATCCGTCCTCAAAAGAAA 1207
Db 434 laArgAspCysSerArgSerArgSerProTyrLysAlaArg---AspArgSerArgSerA 453
QY 1208 GGCTCCGAAGAGAGGAGATATCAC-----CAACCAATCAACTT---GAGA 1252
Db 453 rgSerProAsnLysAlaArgAspHisSerArgSerArgSerProAsnLysAlaArgAspA 473
QY 1253 GAAGGCGAGCCGATCTTTCTAACAACCTTGGGAAGTTATTGAGGTGAAGCCAGACAAG 1312
Db 473 rgSerArgSerArgSer-----ProSerLysGluArgA 484
QY 1313 AAGAACCCCGAGCTTCAGACACTGACATGATCCTCCTGTGTAGAGATCAAGAAGGA 1372
Db 484 sphHisSerGlnLeuGlySerProSerLysGlu-----ArgAspHisArgArgS 500
QY 1373 GCTTTGATGCTCCACACTTCAA-----CTCAAAGGCC 1405
Db 500 er---ArgSerProSerLysGluArgGlnCysArgGlnSerArgSerSerLysGluA 519
QY 1406 ATGTTATCGT-----CGTCGTCAACAAGGAACT 1435
Db 519 rgAspHisArgArgSerArgSerProSerLysGluArgGlnArgArgGlnSerArg-Ser 538
QY 1436 GGAACCTTCAACTCGTGTGTAAGAAAAGAGACAACAGAGGGGACGGCGGGAAGAA 1495
Db 539 ProAsnLysGlu-----ArgAspArgSerGlnSerArgSerProSerGluGlu 554
QY 1496 GAGGAGGACGAGACAGAAAGAGAGGAGGAAGTAACAGAGAGGTGCGTAGGTACACAGCG 1555
Db 555 ArgGluHisArgGlnSerArgSerProSerLysGluArgAspArgArgArgTrpArgSer 574
QY 1556 AGGTTCAAGAGAGCGCATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCT 1615
Db 575 ProSerLysGluArgGlu-----ArgArgGlnSerArgSer 586
QY 1616 TCCTCCGAACCTCCATCTGCTTGGCTTCGGTATCAACGCTGAAAACAACACACAGAACTTTC 1675
Db 587 SerSerGlu-----GluArgAspHisSerArgSer 596
QY 1676 CTTGACAGGTGATAGACAAATGTGATAGACCAGATAGAGAAAGCAAGGAGGATTTAGCA 1735
Db 597 ArgSerProAsnLysGlnSerGlyTyrSerArgProArgAlaSerSerLysGluLysAla 616
QY 1736 TTCCTGGGTGCGGTGAACAAGTTGAGAAGCTCATCAAAAACCAAGGAAATCTCATT 1795
Db 617 HisSerArgSerArgThrProSerLysGluGlyAsnHisSerGlnSerArgThrSerSer 636
QY 1796 GTGAGTGTCTGCTCCTCAATCTCAATCTCCGTCGTCCTCTGAG---AAAGAGTCT 1852
Db 637 LysGluSerAspPro---SerGlnSerThrValProArgSerProAspTrpLysArgSer 655
QY 1853 CTTGAGAAAGAGGATCAAGAGGAGGAAACCAA 1885
Db 656 ProThrArgThrSerSerLeuSerGlnAsnArg 666
```

Search completed: May 16, 2006, 00:04:26
Job time : 92.9525 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 21:46:12 ; Search time 13.6006 Seconds
(without alignments)
4312.597 Million cell updates/sec

Title: US-10-728-323-1
Perfect score: 3586
Sequence: 1 aataacatataattcatc.....cgttgtggtgtttcttccc 2032

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapcp 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p model -DEV=xlp
-Q=/abs/ABSSWEB_spool/US10728323/runat_15052006_172134_22413/app_query.fasta_1
-DB=PIR -OFMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=us10728323 @CN 1 1 92 @runat_15052006_172134_22413 -NCPU=6 -ICPU=3
-NO MAP -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	1282	35.8	605 2 S20007 beta-conglycinin a
2	1281	35.7	605 1 FWSYBA beta-conglycinin a
3	1267	35.3	571 2 S00566 convicillin precurs
4	1251	34.9	639 2 B24810 beta-conglycinin a
5	1178.5	32.9	459 2 S08505 vicilin - garden p
6	1174.5	32.8	433 2 S00567 vicilin precursor
7	1156	32.2	439 1 FWSYCB beta-conglycinin b
8	1148	32.0	438 2 S35757 vicilin, 47K - gar
9	1109.5	30.9	463 2 A27288 vicilin precursor
10	1108.5	30.9	463 2 S06309 vicilin precursor
11	1068	29.8	410 1 FWPMBV vicilin B precursor
12	962.5	26.8	445 2 JQ2264 canavalin - jack b
13	959.5	26.8	445 2 S00281 canavalin - sword
14	907.5	25.3	810 2 T44430 protein PV100 [imp

ALIGNMENTS

RESULT 1

S20007
beta-conglycinin alpha chain precursor - soybean
C:Species: Glycine max (soybean)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S20007
R:Bellevue, J.M.; Dickinson, C.D.; Dickinson, L.A.; Nielsen, N.C.
A:Title: Synthesis and assembly of soybean beta-conglycinin in vitro.
A:Reference number: S20007; MUID:92119248; PMID:1731988
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-605 <LEL>
A:Cross-references: UNIPROT:Q94LX2; UNIPARC:UPI00000A0D19
C:Superfamily: glycinin

Alignment Scores:
Pred. No.: 1.4e-81 Length: 605
Score: 1282.00 Matches: 285
Percent Similarity: 59.4% Conservative: 108
Best Local Similarity: 43.1% Mismatches: 173
Query Match: 35.8% Indels: 96
DB: 2 Gaps: 18

US-10-728-323-1 (1-2032) x S20007 (1-605)

QY 50 ATGAGAGGAGGGGTTTCTCCACTGATCTGTGTTAGGATCTTCTCTGCTTCAGTT 109
Db 2 MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyValPheLeuAlaSerVal 19
QY 110 TTGCAACGCATGCCAAGTCACCTTACCAAGAGAAACAGAACCCCTGCGCCAG 169
Db 20 SerValSerPheGlyIleAla-----TyrTrpGluLys---GluAsnProLysHisAsn 36
QY 170 AGTGCCCTCAGAGTGTCAACAGGACCGGATGACTTGAACGAAAGCATGCGAGTCT 229
Db 37 LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 56

alpha-globulin typ
phaseolin type alp
vicilin precursor
alpha-phaseolin pr
phaseolin beta cha
alpha-globulin typ
alpha-globulin B p
vicilin A precursor
vicilin-like stora
globulin-1S, GLB1S
globulin Beg1 prec
beta-conglycinin a
globulin-2 precurs
62K sucrose-bindin
vicilin-like stora
7S storage protein
convicillin (clone
vicilin-like stora
vicilin - Zamia fu
convicillin precurs
phaseolin - kidney
probable major pro
vicilin, 14K compo
globulin1 - maize
hypothetical prote
probable seed stor
hypothetical prote
globulin-like prot
beta-globulin B pr

A; Cross-references: UNIPARC:UPI00001745FC; UNIPARC:UPI00001745FD; UNIPARC:UPI00001745FE
 A; Experimental source: seed
 C; Superfamily: glycoprotein
 C; Keywords: glycoprotein; seed; storage protein
 F; 1-22/Domain: signal sequence #status predicted <SIG>
 F; 23-62/Domain: propeptide #status predicted <PRO>
 F; 63-605/Product: beta-conglycinin alpha chain #status predicted <MAT>
 F; 261,517/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
 Pred. No.: 1,65e-81 Length: 605
 Score: 1281.00 Matches: 285
 Percent Similarity: 59.2% Conservative: 107
 Best Local Similarity: 43.1% Mismatches: 174
 Query Match: 35.7% Indels: 96
 DB: 1 Gaps: 18

US-10-728-323-1 (1-2032) x FMSYBA (1-605)

QY 50 ATGAGAGGAGGAGGTTCTCCACTGATGCTGTGCTAGGATCCTGTCTCGCTTCAGTT 109
 DB 2 MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyLeuValPheLeuAlaSerVal 19
 QY 110 TGTGCAACGATGCCAAGTATCACCCTTACAGAGAAACAGAGAACCCCTGCGCCAG 169
 DB 20 SerValSerPheGlyIleAla-----TyrTrpGluLys---GluAsnProLysHisAsn 36
 QY 170 AGGTGCTCCAGAGTGTCAACAGGACCGGATGCTTGAAGCAAAAGCGCATGCGAGTCT 229
 DB 37 LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 56
 QY 230 CGCTGCACCAAGCTCGAGTATGAT-----CCTGCTGT 262
 DB 57 ArgCysAsnLeuLeuLysValGluLysGluGluCysGluGluGlyLeuProArg--- 75
 QY 263 GTCTATGATCTCGAGGACACACTGGCACCAACCAACGTTCCCTCCAGGGAGCGG 322
 DB 76 -----ProArgProArg-----ProGlnHisProGlu 84
 QY 323 ACACGTGGCGCCACCCGAGACTACGATGATGAC----- 358
 DB 85 ArgGluProGlnProGlnProGlnGluLysGluGluAspGluAspGluProArgProle 104
 QY 359 -----CGCGCTCAACCCCGAAGAGAG-----GAAGGAGCCGA 391
 DB 105 ProPheProArgProGlnProArgGlnGluGluHisGluGluArgGluGlnGlu 124
 QY 392 TGG-----GGACCAAGTGGACCGAGGAGCGTGAAGAGAGAA 430
 DB 125 TrpProArgLysGluGluLysArgGlyGluLysGlySerGluGluGluAspGlu 144
 QY 431 GACTGGAGACCAAGAGAGAT-----TGGAGCGGACCAAGTCAATCAGCAGCCACGG 484
 DB 145 AspGluGluGlnAspGluArgGlnPheProPheProArgProProHisGlnLysGlu 164
 QY 485 AAAATAAGGCCGCAAGAGAGAAAGAGAGCAAGAGTGGGGAACACCAAGTACGATG 544
 DB 165 ArgAsnGluGluAspGluAspGluGluGlnArgGluSerGluSerGluAsp 184
 QY 545 AGGGAA-----GAAACATCTCGGAACACCTTTCTACTTCCCGTCAAGCGGTTTACG 598
 DB 185 SerGluLeuArgArgHisLysAsnLysAsnProPheLeuPheGlySerAsnArgPheGlu 204
 QY 599 ACCCGTACGGGACCAACCGTAGGATCCGGTCTCGAGAGTTTACCAAGAGTCA 658
 DB 205 ThrLeuPheLysAsnGlnTyrGlyArgIleArgValLeuGlnArgPheAsnGlnArgSer 224
 QY 659 AGGCAGTTTCAGAACTCCAGAAATACCGTATTGTGCAGATCGAGGCCAAACCTAACACT 718
 DB 225 ProGlnLeuGlnAsnLeuArgAspTyrArgIleLeuGluPheAsnSerLysProAsnThr 244
 QY 719 CTTGTTCTTCCCAAGCACCTGATGCTGATAACCTCTTGTATCCAGCAAGGGCAAGCC 778

DB 245 LeuLeuLeuProAsnHisAlaAspAlaAspTyrLeuIleValIleLeuAsnGlyThrAla 264
 QY 779 ACCGTGACCGTAGCAAAATGCGCAATAACAGAAAGAGCTTTAATCTTGACGAGGCCATGCA 838
 DB 265 IleLeuSerLeuValAsnAsnAspArgAspSerTyrArgLeuGlnSerGlyAspAla 284
 QY 839 CTCGAAATCCCATCCGGTTTCATTTCTTACATCTTTGAACCGCCATGACACCAAGAACCTC 898
 DB 285 LeuArgValProSerGlyThrThrTyrValValAsnProAspAsnAsnGluAsnLeu 304
 QY 899 AGAGTAGTAAATCTCCATGCCGCTTAACACACCCGCGCAGTTTGAGGATTTCTTCCCG 958
 DB 305 ArgLeuIleThrLeuAlaIleProValAsnLysProGlyArgPheGluSerPheLeu 324
 QY 959 GCGAGCAGCCGAGACCAATCATCTTCTGAGGCTTTCAGCAGAAATACGTTGAGGCGC 1018
 DB 325 SerSerThrGluAlaGlnSerTyrLeuGlnGlyPheSerArgAsnIleLeuGluAla 344
 QY 1019 GCCTTCAATCGGAATTCATGATACGAGGAGGCTGCTTTAGAGAAATCGACGAGT 1078
 DB 345 SerTyrAspThrLysPheGluGluIleAsnLysValLeuPheSerArgGluGly--- 363
 QY 1079 GAGCAAGAGGAGAGGCGCAGGCGATGCGAGTACTCGAGTAGTGAGAACATCAAGCA 1138
 DB 364 -----GlnGlnGlnGlyGluArgLeu-----GlnGluSer 374
 QY 1139 GTGATAGTCAAGTGTCAAAGGAGCAGCTTGAAGAACTTACTTAAGCACGCTAAATCCGTC 1198
 DB 375 ValIleValGluIleSerLysGlnIleArgAlaLeuSerLysArgAlaLysSerSer 394
 QY 1199 TCAAGAAAGGCTCCGAAGAGAGGAGATATACCAACCCCAATCAACTTTGAGAGAGGC 1258
 DB 395 SerArgLysThrIleSerSerGlu-----AspLysProPheAsnLeuArgSerArg 411
 QY 1259 GAGCCGATCTTTCTAAACAACTTTGCGAAGTTATTTGAGGTGAACCCAGACAGAAAGAAC 1318
 DB 412 AspProIleTyrSerAsnLysLeuGlyLysPheGluIleThrProGlu---LysAsn 430
 QY 1319 CCCAGCTTCAGGACCTGACATGATGCTCACTGCTGTAGATCAAAAGAGGAGCTTGT 1378
 DB 431 ProGlnLeuArgAspLeuAspIlePheLeuSerIleValAspMetAsnGluGlyAlaLeu 450
 QY 1379 ATGCTCCACACTTCAACTCAAAGGCCATGTTATCGTCTGCTCAACAAAGAACTGGA 1438
 DB 451 LeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGlyAspAla 470
 QY 1439 AACCTTGAATCTCGCTGCTGAAGAAAGAGCAACAGAGGGGACGCGGAGAGAGAG 1498
 DB 471 AsnIleGluLeuValGlyLeu---LysGluGlnGlnGlnGlnGlnGlnGlnGln 489
 QY 1499 GAGGACGAAGACGAAGAGAGGAGGAAGTAAACAGAGAGGTCGTAGGTACACAGCAGG 1558
 DB 490 Pro-----LeuGluValArgLysTyrArgAlaGlu 499
 QY 1559 TTGAAGAGAGCGATGTTCTCATCGCAGCTCATCCAGTAGCCATCAACGCTTCC 1618
 DB 500 LeuSerGluGlnAspIlePheValIleProAlaGlyTyrProValValValAsnAlaThr 519
 QY 1619 TCCGACCTCCATCTGCTTGGCTTCCGATCAACGCTGAAACCAACCAAGATCTTCTCT 1678
 DB 520 SerAsnLeuAsnPhePheAlaIleGlyIleAsnAlaGluAsnAsnGlnArgAsnPheLeu 539
 QY 1679 GCAGGTGATAAGGACCAATGTGATAGCAGATAGAGAGCAAGCAAGGATTTAGCATTC 1738
 DB 540 AlaGlySerGlnAspAsnValIleSerGlnIleProSerGlnValGlnGlnLeuAlaPhe 559
 QY 1739 CCTGGTCCGGTGAACAAAGTTGAGAGCTCATCAAAAAACAGAGAAATCTCACTTGTG 1798
 DB 560 ProGlySerAlaGlnAlaValGluLysLeuLeuLysAsnGlnArgGluSerTyrPheVal 579
 QY 1799 AGTGTCTCTCTCAATCTCAATCTCTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1858
 DB 580 AspAlaGlnProLysLys----- 585


```
QY 1505 GAAGACCAAGAGGAGGAGTAAACAGAGAGGTCGCTAGGTACACAGCGAGGTGAAG 1564
||||| :|||
Db 450 GluAspArgLysGluArg-----AsnAsnGluValGlnArgTyrGluAlaArgLeuSer 467
QY 1565 GAAGGCGATGTTCATCATGCGAGCGAGTCATCCAGTAGCCATCAACGCTTCCTCCGAA 1624
||||| :|||
Db 468 ProGlyAspValIleIleProAlaGlyHisProValAlaIleSerAlaSerSerAsn 487
QY 1625 CTCCTCTCTGCTGGCTTCATCAGCTCAACGCTGAACCAACCAACAGATCTTCCTTCAGGT 1684
||||| :|||
Db 488 LeuAsnLeuLeuGlyPheGlyIleAsnAlaLysAsnGlnArgAsnPheLeuSerGly 507
QY 1685 GATAAGCACAATGTGATACACAGATAGAGGAGCAAGCGAAGATTTAGCATTCCTCTGGG 1744
||||| :|||
Db 508 SerAspAspAsnValIleSerGlnIleGluAsnProValLysGluLeuThrPheProGly 527
QY 1745 TCGGGTGAACAAGTTGAGAGCTCATCAAAAACCAAGAGGAATCTCACTTTGTGAGTGTCT 1804
||||| :|||
Db 528 SerSerGlnGluValAsnArgLeuIleLysAsnGlnLysGlnSerHisPheAlaSerAla 547
QY 1805 CGTCTCTCAATCTCAATCTCAATCTCGTCTCTCTCTGAGAAAGAGTCTCTCGAAGAG 1864
||||| :|||
Db 548 GluPro-----Glu 550
QY 1865 GATCAAGAGGAGGAAACCAAGGAGGAGGTCCTCTCTCAATTTTGAAGGCTTTT 1924
||||| :|||
Db 551 GlnLysGluGluGluSerGlnArgLysArgSerProLeuSerSerValLeuAspSerPhe 570

RESULT 4
B24810
beta-conglycinin alpha chain precursor - soybean
C:Species: Glycine max (soybean)
C:Date: 07-Mar-1988 #sequence_revision 30-Jun-1991 #text_change 09-Jul-2004
C:Accession: B24810; S16337
R:Doyle, J.J.; Schuler, M.A.; Godette, W.D.; Zenger, V.; Beachy, R.N.; Slightom, J.L.
J. Biol. Chem. 261, 9228-9238, 1986
A:Title: The glycosylated seed storage proteins of Glycine max and Phaseolus vulgaris. S
A:Reference number: A24810; MUID:86250867; PMID:3013679
A:Accession: B24810
A:Molecule type: DNA
A:Residues: 1-639 <DOY>
A:Cross-references: UNIPROT:P11827; UNIPARC:UPI000012B569; GB:M13759; NID:g169928; PIDN:
A>Note: the authors translated the codon GGT for residue 352 as Glu
R:Schuler, M.A.; Schmitt, E.S.; Beachy, R.N.
Nucleic Acids Res. 10, 8225-8244, 1982
A:Title: Closely related families of genes code for the alpha and alpha' subunits of the
A:Reference number: S16337; MUID:83143288; PMID:6298713
A:Accession: S16337
A:Molecule type: DNA
A:Residues: 361-639 <SCH>
A:Cross-references: UNIPARC:UPI0000177DF9
C:Genetics:
A:Introns: 278/1; 355/3; 382/3; 481/3; 575/1
C:Superfamily: glycinin
C:Keywords: seed, storage protein

Alignment Scores:
Pred. No.: 2,07e-79 Length: 639
Score: 1251.00 Matches: 283
Percent Similarity: 58.7% Conservative: 120
Best Local Similarity: 41.3% Mismatches: 173
Query Match: 34.9% Indels: 110
DB: 2 Gaps: 21

US-10-728-323-1 (1-2032) x B24810 (1-639)
QY 50 ATGAGAGGAGGAGGTTTCTCCACTGATGCTGTTCAGGATCTTGTCTCTGCTTCAGTT 109
Db 2 MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyValValPheLeuAlaSerVal 19
QY 110 TCTGCAACGCATGCCCAAGTCATCACCTTACCAGAGAAACAGAGAACCCCTCGCGCCAG 169
||||| :|||
```

```
Db 20 SerValSerPheGlyIleAla-----TyrTrpGluLys-----GlnAsnProSerHisAsn 36
QY 170 AGGTCCCTCCAGAGTTGTCAACAGGAACCGATGACTTTGAAGCAAAAGCATGCCAGTCT 229
||||| :|||
Db 37 LysCysLeuArgSerCysAsnSerGlnLysAspSerTyrArgAsnGlnAlaCysHisAla 56
QY 230 CGCTGCACC-----AAGCTCGAGTATGATCCTCTCTGT-----CTCTATGAT 271
||||| :|||
Db 57 ArgCysAsnLeuLeuLysValGluGluGluGluCysGluGluGlyGlnIleProArg 76
QY 272 CTTCTGA-----GGACACACTGGCACCAACCAACGCTTCCTCCAGGAGGAGGACACGT 328
||||| :|||
Db 77 ProArgProGlnHisProGluArgGluGlnGlnHisGlyGluLysGluAspGlu 96
QY 329 GSCGCCCAACCGGAGACTACGATGATGACGCG-----CGTCAACCCCGAAGAGAGAA--- 382
||||| :|||
Db 97 GlyGluGlnProArgProPheProArgProGlnProHisGlnGluGlu 116
QY 383 -----GGAGCGCCGATGG 394
Db 117 HisGluGlnLysGluGluHisGluTrpHisArgLysGluGluLysHisGlyGlyLysGly 136
QY 395 GGA-----CCAGCTGGACCGAGGAGCGTGA 421
Db 137 SerGluGluGluGlnAspGluArgGluHisProArgProHisGlnProHisGlnLysGlu 156
QY 422 AGAGAAGAA---GACTGGAGACAACCAAGAGAAGATTGGAGGCGCAAGTCAATCAGCAG 478
||||| :|||
Db 157 GluGluLysHisGluTrpGlnHisLysGlnGluLysHisGlnGlyLysGluSerGluGlu 176
QY 479 CCACGGAATAAGGCCCGCAAGAGAGAGAAAGAGAACAAAGAG-----TGGGGAACACACAGT 535
||||| :|||
Db 177 GluGluGluAspGlnAspGluGluGlnAspLysGluSerGlnGluSerGluGly 196
QY 536 AGCCATGTGAGGAGAAACAATCTCGAAC-----AACCTTTCTACTTCCGTC 586
||||| :|||
Db 197 SerGluSerGlnArgGluProArgHisLysAsnLysAsnProPheHisPheAsnSer 216
QY 587 AGGCGGTTTACACCGCTAGCGGAACCAACCGTAGGATCCGGTCTCTCAGAGGTTT 646
||||| :|||
Db 217 LysArgPheGlnThrLeuPheLysAsnGlnTyrGlyHisValArgValLeuGlnArgPhe 236
QY 647 GACCAAAAGTCAAGGACGTTTTCAGAAATCTCAGAAATCACCGTATTGTGCAGATCGAGCC 706
||||| :|||
Db 237 AsnLysArgSerGlnGlnLeuGlnAsnLeuArgAspTyrArgIleLeuGluPheAsnSer 256
QY 707 AAACCTAACACTTTGTTCTTCCCAAGCAGCGTGTATGTGTATACATCCTTTTATCCAG 766
||||| :|||
Db 257 LysProAsnThrLeuLeuLeuProHisHisAlaAspAlaAspTyrLeuIleValIleLeu 276
QY 767 CAAGGCAAGCCACCGTCACCGTAGCAATGGCAATACAGAAAGAGCTTTAATCTTGAC 826
||||| :|||
Db 277 AsnGlyThrAlaIleLeuThrLeuValAsnAsnAspArgAspSerTyrAsnLeuGln 296
QY 827 GAGGCGCATGCACTCAGAAATCCATCCGTTTCATTCTCATCTTGAAC----- 877
||||| :|||
Db 297 SerGlyAspAlaLeuArgValProAlaGlyThrPheTyrValValAsnProAspAsn 316
QY 878 -----CGCATGACAAAC----- 889
Db 317 AspGluAsnLeuArgMetIleAlaGlyThrThrPheTyrValValAsnProAspAsnAsp 336
QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCCCGTTTAAACACACCCCGGAGTTTCAGGAT 949
||||| :|||
Db 337 GluAsnLeuArgMetIleThrLeuAlaIleProValAsnLysProGlyArgPheGluSer 356
QY 950 TTCTTCCCGGAGCAGCGGAGACCAATCATCTACTTTCAGGGCTTCAGCAGGAATACG 1009
||||| :|||
Db 357 PhePheLeuSerSerThrGlnAlaGlnSerTyrLeuGlnGlyPheSerLysAsnIle 376
QY 1010 TTGAGGCGGCTTCAATCGCGAATTCATAGATACGAGGAGGCTGCTGTAGAGAGAAAT 1069
||||| :|||
Db 377 LeuGluAlaSerTyrAspThrLysPheGluGluIleAsnLysValLeuPheGlyArgGlu 396
```

QY	1070	GCAGGAGTTCAGCAA---GAGGAGAGAGCGCAGCGGATGGAGTACTCGGAGTAGTGAG	1126
Db	397	:	407
QY	1127	AACAATGAGGAGTGATAGTCAAGTGTCAAAGGAGCAGCTTGAAGAACTTACTAAGCAC	1186
Db	408	-----GluSerValIleValGluIleSerLysGlnIleArgGluLeuSerLysHis	425
QY	1187	GCTAAATCCGCTCTCAAGAAAGGCTCCGAAGAGAGGAGATATCAACCAACCAATCAAC	1246
Db	426	AlaLysSerSerSerArgLysThrIleSerSerGlu-----AspLysProPheAsn	442
QY	1247	TTGAGAGAGGCGAGCCCGATCTTTCTTAACAATCTTGGGAAGTATTGAGGTGAAGCCA	1306
Db	443	LeuGlySerArgAspProIleTyrSerAsnLysLeuGlyLysLeuPheGluIle-----	460
QY	1307	GACAAGAAGAACCCAGCTTCAGGACCTGGACATGATGTCACCTGTGTAGAGATCAAA	1366
Db	461	ThrGlnArgAsnProGlnLeuArgAspLeuAspValPheLeuSerValValAspMetAsn	480
QY	1367	GAAGGAGCTTTGATCTCCACACTTCAACTCAAAGGCCATGGTTATCGTCTGTCACAC	1426
Db	481	GluGlyAlaLeuPheLeuProHisPheAsnSerLysAlaIleValValLeuValIleAsn	500
QY	1427	AAAGGAACCTGGAAACCTTGAACTCGTGGCTGTGAAGAAAGAGCAACAACAGGGGACGG	1486
Db	501	GluGlyGluAlaAsnIleGluLeuValGlyIle---LysGluGlnGlnArgGlnGln	519
QY	1487	CGGAGAGAGAGGAGCGAAGACGAAGAGAGGGAAGTAACAGAGGTGCGTAGG	1546
Db	520	GlnGluGlnPro-----LeuGluValArgLys	529
QY	1547	TACACAGCAGAGGTGAAGGAAGCGATGTCTTCATCGCCAGCAGCTCATCCAGTAGCC	1606
Db	530	TyrArgAlaGluLeuSerGluGlnAspIlePheValIleProAlaGlyTyrProValMet	549
QY	1607	ATCAAGCTTCTCCGAACTCCATCTGCTGGCTTCGGTATCAACGCTGAAGAAACACACAC	1666
Db	550	ValAsnAlaThrSerAspLeuAsnPhePheAlaPheGlyIleAsnAlaGluAsnAsnGln	569
QY	1667	AGAATCTTCCTCCAGTGATAGGACAAATGTATAGACCAGATAGAGAGCAAGCAAG	1726
Db	570	ArgAsnPheLeuAlaGlySerLysAspAsnValIleSerGlnIleProSerGlnValGln	589
QY	1727	GATTTAGCATTCCTCGGCTCGGTGAACAAGTTGAGAAGCTCATCAAAAAACCAAGGAA	1786
Db	590	GluLeuAlaPheProArgSerAlaLysAspIleGluAsnLeuIleLysSerGlnSerGlu	609
QY	1787	TCTCATTGTGTAGTGTCTGCTCTCAATCTCAATCTCAATCTCGTCTCTCTAGAAA	1846
Db	610	SerTyrPheValAspAlaGlnProGln-----	618
QY	1847	GAGTCTCTGAGAAAGAGGATCAAGAGGAGGAAACCAAGGAGGAGGTCCACTCCTT	1906
Db	619	-----GlnLysGluGluGlyAsnLysGlyArgLysGlyProLeuSer	632
QY	1907	TCAATTTTGAAGGCTTTT	1924
Db	633	SerIleLeuArgAlaPhe	638
RESULT 5			
S08505			
vicilin - garden pea			
C;Species: Pisum sativum (garden pea)			
C;Date: 02-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 05-Oct-2004			
C;Accession: S08505			
R;Higgins, T.J.V.; Newbigin, E.J.; Spencer, D.; Llewellyn, D.J.; Craig, S.			
Plant Mol. Biol. 11, 683-695, 1988			
A;Title: The sequence of a pea vicilin gene and its expression in transgenic tobacco pla			
A;Reference number: S08505			
A;Accession: S08505			
A;Status: preliminary			

A:Molecule type: DNA
A:Residues: 1-459 <HIG>
A:Cross-references: UNIPROT:P13918; UNIPARC:UPI0000138277; EMBL:X14076; NID:g20917; PID:
C:Genetics:
A:Introns: 102/1; 160/3; 187/3; 295/3; 400/1
C:Superfamily: vicilin

Alignment Scores:	2,22e-74	Length:	459
Pred. No.:	1178.50	Matches:	223
Score:	71.3%	Conservative:	93
Percent Similarity:	50.3%	Mismatches:	114
Best Local Similarity:	32.9%	Indels:	13
Query Match:	2	Gaps:	4
DB:			

US-10-728-323-1 (1-2032) x S08505 (1-459)

QY	542	GTGAGGGAAGAAACATCTCGGAACAACCCCTTTCTACTTCCCTCAAGCGGTTTACACC	601
Db	26	ValSerSerArgSerAspProGlnAsnProPheIlePheLysSerAsnLysPheGlnThr	45
QY	602	CGCTACGGGAACCAAAACGGTAGGATCCGGGTCTGCAGAGGTTTGACCAAGGTCACAGG	661
Db	46	LeuPheGluAsnGluAsnGlyHisIleArgLeuLeuGlnLysPheAspGlnArgSerLys	65
QY	662	CAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTAACACTCTT	721
Db	66	IlePheGluAsnLeuGlnAsnTyrArgLeuLeuGluTyrLysSerLysProHisThrIle	85
QY	722	GTCTTCCCAAGCAGCTCATGCTGATAACATCTCTTGTATTATCCAGCAAGCGCAAGCCACC	781
Db	86	PheLeuProGlnHisThrAspAlaAspTyrIleLeuValValLeuSerGlyLysAlaIle	105
QY	782	GTACCGTACCAAAATGGCAATAACAGAAAGACTTTAAATCTTCACGAGGGCCATGCACATC	841
Db	106	LeuThrValLeuLysProAspArgAsnSerPheAsnLeuGluArgGlyAspThrIle	125
QY	842	AGAAATCCATCCGGTTTCATTCTTCATCATCTTGACCCGCCATCACACCAAGACCTCAGA	901
Db	126	LysLeuProAlaGlyThrIleAlaTyrLeuValAsnArgAspAsnGluGluLeuArg	145
QY	902	GTAGCTAAAATCTCCATGCCCGCTTAACACACCCGCCAGTTTGAGATTCTTCCCGGCG	961
Db	146	ValLeuAspLeuAlaIleProValAsnArgProGlyGlnLeuGlnSerPheLeuLeuSer	165
QY	962	AGCAGCCGAGACAAATCATCTTTCAGGGCTTCAGCAGGAATACGTTGGAGGCGGCC	1021
Db	166	GlyAsnGlnAsnGlnGlnAsnTyrLeuSerGlyPheSerLysAsnIleLeuGluAlaSer	185
QY	1022	TTCAATCCGGAATTCATAGATACGAGAGGTGCTGTTAGAGAGAAATGCAGAGGTGAG	1081
Db	186	PheAsnThrAspTyrGluGluIleGluLysValLeuLeuGluGluHisGluLysGluThr	205
QY	1082	CAAGAGGAGAGGGCAGAGCGGATCGAGTACTCGGAGTAGTGAGAACAAATGAAGAGTG	1141
Db	206	GlnHisArgArgSerLeuLys-----AspLysArgGlnGlnSerGlnGluAsnVal	223
QY	1142	ATAGTCAAAGTGTCAAAGAGCAGCTTGAAGAACTTACTTAAGCAGCGTAAATCCGTCTCA	1201
Db	224	IleValLysLeuSerArgGlyGlnIleGluLeuSerLysAsnAlaLysSerThrSer	243
QY	1202	AAGAAAGGCTCCGAAGAAGGAGGATATCACCAACCCCAATCACTTGAGAGAAGCGCAG	1261
Db	244	LysLysSerValSerSerGluSerGlu-----ProPheAsnLeuArgSerArgGly	260
QY	1262	CCCGATCTTCTAACCAACTTTGGGAAAGTTATTGAGGTGAAGCCAGACAAGAAAGACCCC	1321
Db	261	ProIleTyrSerAsnGluPheGlyLysPhePheGluIleThrProGlu---LysAsnPro	279
QY	1322	CAGCTTCAGGACCTGGACATGATGCTTCACCTGTGTAGATCAAGAGGAGCTTTGATG	1381
Db	280	GlnLeuGlnAspLeuAspIlePheValAsnSerValGluIleLysGluGlySerLeuLeu	299

```
QY 1382 CTCGCCACATTCACAAAGGCCATGGTTATCGTCGTCTCAACAAAGAACTGGAAC 1441
Db 300 LeuProHisTyrAsnSerArgAlaIleValIleValThrValAsnGluGlyValGlyAsp 319
QY 1442 CTTGAACTCGTCTGTAGAAAGAGCAACACAGAGGGGACGGCGGAAGAAGAGAG 1501
Db 320 PheGluLeuValGlyGlnArgAsnGluAsnGlnGlnGluGlnArgGlyGluAspGlu 339
QY 1502 GACGAAGACGAAGAAGAGAGGGAAGTAACAGAGAGGTGCGTAGGTACACAGCGGTTG 1561
Db 340 GluGluGluGlnGlyGluGluGluIleAsnLysGlnValGlnAsnTyrLysAlaLysLeu 359
QY 1562 AAGGAAGGCGATGTTTCATCATGCGCAGCAGCTCATCCAGTAGCCATCAACGCTTCCTCC 1621
Db 360 SerSerGlyAspValPheValIleProAlaGlyHisProValAlaValLysAlaSerSer 379
QY 1622 GAATCCATCTGCTCGCTTCGGTATCAACGCTGAAACCAACACACAGATCTTCCTTGA 1681
Db 380 AsnLeuAspLeuLeuGlyPheGlyIleAsnAlaGluAsnGlnArgAsnPheLeuAla 399
QY 1682 GGTGATAAGGACAATGTGATAGACCAAGAGCAAGCAAGGATTTAGCATTTCCCT 1741
Db 400 GlyAspGluAspAsnValIleSerGlnIleGlnArgProValLysGluLeuAlaPhePro 419
QY 1742 GGTGCGGTGACAAAGTTCAGAGCTCATCAAAACCAAGAGGAATCTCACTTGTGAGT 1801
Db 420 GlySerAlaGlnGluValAspArgIleLeuGluAsnGlnLysGlnSerHisPheAlaAsp 439
QY 1802 GCTCGTCTCAATCTCAATCTCAATCTCGCTCTCTGAGAGAGAGTCTCTGAGAA 1861
Db 440 AlaGlnProGlnArg-----GluArgGlySerArgGluThr 452
QY 1862 GAGGATCAA 1870
Db 453 ArgAspArg 455

RESULT 6
S00567
vicilin precursor (clone pDUB9) - garden pea (fragment)
C:Species: Pisum sativum [garden pea]
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 05-Oct-2004
C:Accession: S00567
R:Watson, M.D.; Lambert, N.; Delauney, A.; Yarwood, J.N.; Croy, R.R.D.; Gatehouse, J.A.;
Biochem. J. 251, 857-864, 1988
A:Title: Isolation and expression of a pea vicilin cDNA in the yeast Saccharomyces cerev
A:Reference number: S00567; MUID:88326226; PMID:3046604
A:Accession: S00567
A:Molecule type: mRNA
A:Residues: 1-433 <WAT>
A:Cross-references: UNIPROT:PI3918; UNIPARC:UPI000016DF3C; EMBL:Y00722; NID:g20915; PIDN
C:Superfamily: vicilin
F:1-2/Domain: signal sequence (fragment) #status predicted <SIG>
F:3-433/Product: vicilin #status predicted <MAT>

Alignment Scores:
Pred. No.: 4,17e-74 Length: 433
Score: 1174.50 Matches: 221
Percent Similarity: 72.2% Conservative: 93
Best Local Similarity: 50.8% Mismatches: 108
Query Match: 32.8% Indels: 13
DB: 2 Gaps: 4

US-10-728-323-1 (1-2032) x S00567 (1-433)

QY 566 AACCTTTTCTACTTCCGTCACAGGCGGTTTAGACCCGCTACGGNACCAAAACGGTAGG 625
Db 8 AsnProPheIlePheLysSerAsnLysPheGlnThrLeuPheGluAsnGluAsnGlyHis 27
QY 626 ATCCGGGTCTCGAGAGGTTTGACAAAGGTCAAGGCACTTTTCAGATCTCCAGATATCAC 685
Db 28 IleArgLeuLeuGlnLysPheAspGlnArgSerLysIlePheGluAsnGlnAsnTyr 47
QY 686 CGTATTGTGCAGATCGAGGCCAAACCTAACACTCTTGTGTTCTTCCCAAGCACGCTGATGCT 745
```

```
Db 48 ArgLeuLeuGluTyrLysSerLysProHisThrIlePheLeuProGlnHisThrAspAla 67
QY 746 GATACATCTCTGTTATCCAGCAAGGCAAGCCACCGTACCGTAGCAAAATGCCAATAAC 805
Db 68 AspTyrIleLeuValValLeuSerGlyLysAlaIleLeuThrValLeuLysProAspAsp 87
QY 806 AGAAAGAGCTTTAATCTTTGACGAGGCGCATGCACTCAGAAATCCCATCCGTTTCATTTCC 865
Db 88 ArgAsnSerPheAsnLeuGluArgGlyAspThrIleLysLeuProAlaGlyThrIleAla 107
QY 866 TACATCTTGAACCGCCATGACCAACAGAACCTCAGAGTAGCTAAATACTCCATGCCGCTT 925
Db 108 TyrLeuValAsnArgAspAsnGluGluLeuArgValLeuAspLeuAlaIleProVal 127
QY 926 AACACACCGCGCAGTTTGGAGATTCTTCCGCGGAGCAGCGCAGACCAATCATCTCTAC 985
Db 128 AsnArgProGlyGlnLeuGlnSerPheLeuLeuSerGlyAsnGlnAsnGlnGlnAsnTyr 147
QY 986 TTGCAAGGCTTTCAGCAGGAATACGTTGGAGCGCGCTTCAATTCGGAATTTCAATGAGATA 1045
Db 148 LeuSerGlyPheSerLysAsnIleLeuGluAlaSerPheAsnThrAspTyrGluGluIle 167
QY 1046 CGGAGGTGCTGTTTGAAGAGAAATCCAGAGGTGAGCAAGAGGAGGAGGCGCAGAGCGCA 1105
Db 168 GluLysValLeuLeuGluGluHisGluLysGluThrGlnHisArgArgSerLeuLys--- 186
QY 1106 TGGAGTACTCGAGTAGTGAGAAACAATGAAGGAGTGATGTCAAAGTGTCAAAGGAGCAC 1165
Db 187 ---AspLysArgGlnGlnSerGlnGluAsnValIleValLysLeuSerArgGlyGln 205
QY 1166 GTTGAAGAACTTACTAAGCACGCTAAATCCGCTCTCAAAAGAAAGGCTCCGAAAGAGGGA 1225
Db 206 IleGluGluLeuSerLysAsnAlaLysSerThrSerLysLysSerValSerSerGluSer 225
QY 1226 GATATCAACACCGCAATCACTTGAGAGAGCGGAGCGCGATCTTCTTAACACTTTGGG 1285
Db 226 Glu-----ProPheAsnLeuArgSerArgGlyProIleTyrSerAsnGluPheGly 242
QY 1286 AAGTTATTGAGTGTAAGCAAGCAAGAAAGACCCCGAGCTTCAGGACCTGCACATGATG 1345
Db 243 LysPhePheGluIleThrProGlu---LysAsnProGlnLeuGlnAspLeuAspIlePhe 261
QY 1346 CTCACCTGTGTAGAGATCAAAAGAGAGGAGCTTTGATGCTCCCACTTCAACTCAAGGCGC 1405
Db 262 ValAsnSerValGluIleLysGluGlySerLeuLeuLeuProHisTyrAsnSerArgAla 281
QY 1406 ATGTTTATCGTCTGCTCAACAAAGGAACCTTGAACCTTGAACCTGCTGCTGTAGAAAA 1465
Db 282 IleValIleValThrValAsnGluGlyLysGlyAspPheGluLeuValGlyGlnArgAsn 301
QY 1466 GAGCAACAACAGAGGGGACGGCGGAAGAAGAGGAGGACGAAGACGAAGAGAGGAGGGA 1525
Db 302 GluAsnGlnGlnGlnArgLysGluAspGluGluGluGluGluGluGluGlu 321
QY 1526 AGTAACAGAGAGGTGCTAGGTATACACAGCGAGGTTTGAAGGAGGCGATGTGTTTCATG 1585
Db 322 IleAsnLysGlnValGlnAsnTyrLysAlaLysLeuSerSerGlyAspValPheValIle 341
QY 1586 CCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCGAACTCCATCTGCTTGGCTTCGGT 1645
Db 342 ProAlaGlyHisProValAlaLeuLysAlaSerSerAsnLeuAspLeuGlyPheGly 361
QY 1646 ATCAACGCTGAAACAACACACAGATCTTCTTGAGGTGATGAAGCAAACTGTGATAGAC 1705
Db 362 IleAsnAlaGluAsnAsnGlnArgAsnPheLeuAlaGlyAspGluAspAsnValIleSer 381
QY 1706 CAGATAGAGAAGCAAGCGAAGATTTAGCATTTCCCTGGGTCCGGTGAACAAGTTGAGAAG 1765
Db 382 GlnValGlnArgProValLysGluLeuAlaPheProGlySerAlaGlnGluValAspArg 401
QY 1766 CTCATCAAAAACCAAGAGGAATCTCACTTTGTGAGTGTCTGCTCTCAATCTCAATCTCAA 1825
Db 1825
```


A:Cross-references: UNIPROT:Q43626; UNIPARC:UPI000009F8DC; EMBL:X67429; NID:g297169; PID C:Superfamily: vicilin

Alignment Scores:
Pred. No.: 2,96e-72 Length: 438
Score: 1148.00 Matches: 222
Percent Similarity: 73.6% Conservative: 90
Best Local Similarity: 52.4% Mismatches: 100
Query Match: 32.0% Indels: 12
DB: Gaps: 5

US-10-728-323-1 (1-2032) x S35757 (1-438)

```
QY 542 GTGAGGAAAGAACTCTCGGAACACCCCTTTCTACTTCCCGTCAAGGGGTTTAGCAC 601
|||
Db 22 ValSerSerArgSerAspGlnGluAsnProPheIlePheLysSerAsnArgPheGlnThr 41
QY 602 CGGTACGGGAACCAAAACCGTAGGATCCGGTCTCGCAGAGGTTTGACCAAGAGTCAAG 661
|||
Db 42 LeuTyrGluAsnGluAsnGlyHisIleArgLeuLeuGlnLysPheAspLysArgSerLys 61
QY 662 CAGTTTCAGAACTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTAACACTCT 721
|||
Db 62 IlePheGluAsnLeuGlnAsnTyrArgLeuLeuGlnLysSerLysProHisThrLeu 81
QY 722 GTTCTTCCCAAGCACGCTGATGCTGATAATCATCTTGTATTATCCAGCAAGGCCAAGCCAC 781
|||
Db 82 PheLeuProGlnTyrThrAspAlaAspPheIleLeuValValLeuSerGlyLysAlaThr 101
QY 782 GTGACCGTACCAATCGCAATACAGAAAGAGCTTTAATCTTCAGCAGAGGCCATGCATC 841
|||
Db 102 LeuThrValLeuLysSerAsnAspArgAsnSerPheAsnLeuGluArgGlyAspAlaIle 121
QY 842 AGAATCCCATCCGTTTCTTCTACATCTTGAACCGCCAGTCTTTCAGGATTTCTTCCCGGG 901
|||
Db 122 LysLeuProAlaGlyThrIleAlaTyrLeuAlaAsnArgAspAsnGlnAspLeuArg 141
QY 902 GTAGCTPAAATCTCCATGCCGTTTAAACACACCCCGCCAGTCTTTCAGGATTTCTTCCCGGG 961
|||
Db 142 ValLeuAspLeuAlaIleProValAsnLysProGlyGlnLeuGlnSerPheLeuLeuSer 161
QY 962 AGCAGCGAGACCAATCATCTCTTTCAGGGGCTTCAGCAAGAAATAGTTGGAGGCGCC 1021
|||
Db 162 GlyThrGlnAsnGlnProSerLeuLeuSerGlyPheSerLysAsnIleLeuGluAlaIle 181
QY 1022 TTCAATGCGGAATTCAATCAGATACCGAGGGTCTGTTAGAGAGAAATCCAGAGGTGAG 1081
|||
Db 182 PheAsnThrAsnTyrGluGluIleGluLysValLeuLeuGluGlnGln-----Glu 198
QY 1082 CAAGAGGAGAGGCGAGGCGAGTGGAGTACTCTGGAGTAGTGAG---AACAAATGAAGGA 1138
|||
Db 199 GlnGluProGlnHisArgSerLeuLysAspArgArgGlnGluIleAsnGluGluAsn 218
QY 1139 GTGATAGTCAAGTGTCAAGAGACAGCTTGAAGAACTTACTTAAGCACCGCTAAATCCGTC 1198
|||
Db 219 ValIleValLysValSerArgGluGlnIleGluGluLeuSerLysAsnAlaLysSerSer 238
QY 1199 TCAAGAAAGGCTCCGAAAGAGGAGATATACCAACCCCAATCAACTTGTAGAGAGGC 1258
|||
Db 239 SerLysLysSerValSerSerGlu-----SerglyProPheAsnLeuArgSerArg 255
QY 1259 GAGCCGATCTTTCTAAACAACTTTGGGAAGTTATTGGGTGAAGCCAGACAAAGAAAC 1318
|||
Db 256 AsnProIleTyrSerAsnLysPheGlyLysPhePheGluIleThrProGlu---LysAsn 274
QY 1319 CCCAGCTTCAGACCTGACATGATGCTCACCTGTGTAGAGATCAAGAAGAGGCTTTG 1378
|||
Db 275 GlnGlnLeuGlnAspLeuAspIlePheValAsnSerValAspIleLysGluGlySerLeu 294
QY 1379 ATGCTCCCACTTCAACTCAAGGCCATGGTTATCGTCTGCTCAACAAGAGACTGGA 1438
|||
Db 295 LeuLeuProAsnTyrAsnSerArgAlaIleValIleValThrValThrGluGlyLysGly 314
```

```
QY 1439 AACCTTGAACCTCGTGGCTGTAGAAAAAGACCAACACAGAGGCGGCGGAAGAAGAG 1498
|||
Db 315 AspPheGluLeuValGlyGlnArgAsnGluAsnGlnGlyLys-----GluAsn 330
QY 1499 GAGGACGAGAACGAAGAGAGAGGGAAGTAAACAGAGAGGTCGCTAGGTACACGAGGAG 1558
|||
Db 331 AspLysGluGluGlnGluGlnGluThrSerLysGlnValGlnLeuTyrArgAlaLys 350
QY 1559 TTGAAGGAAGGCGATGTTTCATCATCCAGCAGCTCATCCAGTAGTACCATCAACGCTTCC 1618
|||
Db 351 LeuSerProGlyAspValPheValIleProAlaGlyHisProValAlaIleAsnAlaSer 370
QY 1619 TCCGAATCCATCTCTGGCTTCGGTATCAACGCTGAAAAACACACACCAAAATCTTCCTT 1678
|||
Db 371 SerAspLeuAsnLeuIleGlyPheGlyIleAsnAlaGluAsnGlnArgAsnPheLeu 390
QY 1679 GCAGGTGATAGGACAATGTGATACACCATAGATAGAGAACGACCAAGATTTAGCATTC 1738
|||
Db 391 AlaGlyGluGluAspAsnValIleSerGlnValGluArgProValLysGluLeuAlaPhe 410
QY 1739 CQTGGTCCGGTGAAACAAGTTGAGAGCTCATCAAAAACCAAGAAATCTCCTTTGTG 1798
|||
Db 411 ProGlySerSerHisGluValAspArgLeuLeuLysAsnGlnLysGlnSerTyrPheAla 430
QY 1799 AGTGTCTGCTCT 1810
|||
Db 431 AsnAlaGlnPro 434
RESULT 9
A27288
vicilin precursor - fava bean
C:Species: Vicia faba (fava bean)
C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 05-Oct-2004
C:Accession: A27288
R:Basuener, R.; Van Hai, N.; Jung, R.; Saalbach, G.; Muentz, K.
Nucleic Acids Res. 15, 9609, 1987
A:Title: The primary structure of the predominating vicilin storage protein subunit from
A:Reference number: A27288; MUID:88067789; PMID:3684610
A:Accession: A27288
A:Molecule type: mRNA
A:Residues: 1-463 <BAS>
A:Cross-references: UNIPROT:P08438; UNIPARC:UPI000016E014; GB:Y00462; NID:g22052; PIDN:
C:Superfamily: vicilin
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-463/Product: vicilin #status predicted <VCN>
Alignment Scores:
Pred. No.: 1,46e-69 Length: 463
Score: 1109.50 Matches: 218
Percent Similarity: 68.9% Conservative: 97
Best Local Similarity: 47.7% Mismatches: 119
Query Match: 30.9% Indels: 23
DB: Gaps: 4
US-10-728-323-1 (1-2032) x A27288 (1-463)
QY 554 ACATCTCGGAACAACCCCTTTCTACTTCCGTCGTCGTCAGCGGCTTTAGCACCGCTACGGGAC 613
|||
Db 30 SerAspGlnAspAsnProPheValPheGluSerAsnArgPheGlnThrLeuPheGluAsn 49
QY 614 CAAACCGTAGATCCCGGTCCTGCAGAGGTTTGACCAAGGTTTCAAGCGAGTTCAGAAT 673
|||
Db 50 GluAsnGlyHisIleArgLeuLeuGlnLysPheAspGlnHisSerLysLeuLeuGluAsn 69
QY 674 CTCGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTTAACACTCTTCTTCTTCCCAAG 733
|||
Db 70 LeuGlnAsnTyrArgLeuLeuGlnTyrLysSerLysProHisThrIlePheLeuProGln 89
QY 734 CACGCTGATCTGATAACATCTTGTATTATCCAGCAAGGCGCACCGTCGACCGCTAGCA 793
|||
Db 90 GlnThrAspAlaAspPheIleLeuValLeuSerGlyLysAlaIleLeuThrValLeu 109
QY 794 AATGCAATAACAGAAAGAGCTTTAATCTTGACGAGGCCATGCACTCAGAAATCCCATCC 853
```



```
QY 974 CAATCATCTCTTGCAGGCGCTTTCAGCAGGAATACGTTGGAGGCCCGCTTCAATCGGAA 1033
Db 170 GlnProSerIleLeuSerGlyPheSerLysAsnIleLeuGluAlaSerPheAsnThrAsp 189
QY 1034 TTCAATGATAGATCGGAGGTCGTTTGAAGAAATGCGAGGAGTGAAGAGGAGAGAA 1093
Db 190 TyrLysGluIleGluLysValLeuLeuGluGluHisGlyLysGluLysTyrHisArgArg 209
QY 1094 GGCAGAGCGGATGGAGTACTCGGAGTAGTGAGACAATGAGAGGAGTGTAGTCAAGTGG 1153
Db 210 GlyLeuLysAspArgGlnArgGlyGlnGluLysAsn-----ValIleValLysIle 227
QY 1154 TCAAGGAGCAGCGTTGAAGAACTTACTTAAGCAGCGTAAATCCGCTCAAGAAAGCGCTCC 1213
Db 228 SerArgLysGlnIleGluLeuLeuAsnLysAsnAlaLysSerSerLysLysSerThr 247
QY 1214 GAAGAAGAGGAGATATACCAACCCCAATCACTTGAGAGAAGCGAGCCGATCTTTCT 1273
Db 248 SerSerGluSerGlu-----ProPheAsnLeuArgSerArgGluProIleTyrSer 264
QY 1274 AACAACCTTTGGGAAGTTATTGAGGTGAAGCCAGACAAGAAAGACCCAGCTTTCAGGAC 1333
Db 265 AsnLysPheGlyLysPheGluIleThrPro---LysArgAsnProGlnLeuGlnAsp 283
QY 1334 CTGCACATGATGCTACCTGTGTAGAGATCAAAAGAGGAGCTTTGTATGCTCCACACTTC 1393
Db 284 LeuAsnIlePheValAsnTyrValGluIleAsnGluGlySerLeuLeuLeuProHisTyr 303
QY 1394 AACTCAAGGCCATGTTATCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1453
Db 304 AsnSerArgAlaIleValIleValThrValAsnGluGlyLysGlyAspPheGluLeuVal 323
QY 1454 GCTGTAAAGAAAGAGCAACAACAGAGGGGACGCGGAGAGAGAGGAGGAGGAGGAGAA 1513
Db 324 GlyGlnArgAsnGluAsnGlnGlnGlyLeuArgGluGluTyrAspGluGluLysGluGln 343
QY 1514 GAAGAGGAGGAGGAGTAAACAGAGAGGTGCGTAGGTACACAGCGAGGAGGAGGAGGAGG 1573
Db 344 GlyGluGluGluIleArgLysGlnValGlnAsnTyrLysAlaLysLeuSerProGlyAsp 363
QY 1574 GTGTTTCATGATCCAGCAGCTCATCCAGTAGCCATCAACGCTTCCTCCGAACTCCATC 1633
Db 364 ValLeuValIleProAlaGlyTyrProValAlaIleLysAlaSerSerAsnLeuAsnLeu 383
QY 1634 CTGGGCTTCGGTATCAACGCTGAAACCAACAGCAATCTTCTTCAGGTGATAAGGAC 1693
Db 384 ValGlyPheGlyIleAsnAlaGluAsnAsnGlnArgTyrPheLeuAlaGlyGluGluAsp 403
QY 1694 AATGTGATAGACAGATAGAGAGCAAGCAAGGATTTAGCATTCCTCGGTCGGGTGAA 1753
Db 404 AsnValIleSerGlnIleHisLysProValLysGluLeuAlaPheProGlySerAlaGln 423
QY 1754 CAAGTTGAGAGCTCATCAAAACCAAGAGGAATCTCACTTTGTAGTGTCTGCTCTCAA 1813
Db 424 GluValAspThrLeuLeuGluAsnGlnLysGlnSerHisPheAlaAsnAlaGlnProArg 443
QY 1814 TCTCAATCTCAATCTCGTCTCTCTGAGAAAGAGTCTCTCAGAGAAAGAGGATCAAGAG 1873
Db 444 -----GluArgGlu 446
QY 1874 GAGGAAACCAAGAGGAGGAGGTCCTCTTCAATTTTGAAGCTTTT 1924
Db 447 ArgGlySerGlnGluIleLysAspHisLeuTyrSerIleLeuGlySerPhe 463

RESULT 11
FWPMVB
vicilin B precursor - garden pea (fragment)
C:Species: Pisum sativum (garden pea)
C:Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 05-Oct-2004
C:Accession: A03344
R:lycett, G.W.; Delauney, A.J.; Gatehouse, J.A.; Gilroy, J.; Croy, R.R.D.; Boulter, D.
Nucleic Acids Res. 11, 2367-2380, 1983
A:Title: The vicilin gene family of pea (Pisum sativum L.): a complete cDNA coding sequ
```

```
A:Reference number: A93462; MUID:83220791; PMID:6687941
A:Accession: A03344
A:Molecule type: mRNA
A:Residues: 1-410 <LYC>
A:Cross-references: UNIPROT:P02854; UNIPARC:UPI0000138276
A:Experimental source: cv. Feltham First, clones pDUB7 and pDUB4
A:Note: parts of this sequence, including the amino end of the mature protein, were det.
C:Comment: The gene that codes for this protein is part of a multigene family coding fo.
C:Superfamily: vicilin
C:Keywords: glycoprotein
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-410/Product: vicilin type B (fragment) #status predicted <MAT>
F:321-322/Cleavage site: Asn-Asp (unidentified proteinase) #status experimental
F:359/Binding site: carboxylate (Asn) (covalent) #status experimental

Alignment Scores:
Pred. No.: 1,12e-66 Length: 410
Score: 1068.00 Matches: 209
Percent Similarity: 71.6% Conservative: 84
Best Local Similarity: 51.1% Mismatches: 104
Query Match: 29.8% Indels: 12
DB: 1 Gaps: 5

US-10-728-323-1 (1-2032) x FWPMVB (1-410)
QY 542 GTGAGGGAAGAAACATCTCGAACAACCCCTTTCTACTTCCCGTCAAGCGGTTTAGCACC 601
Db 13 ValSerSerArgSerAspGlnGluAsnProPheIlePheLysSerAsnArgPheGlnThr 32
QY 602 CCTACGGGAACCAAAACGGTAGGATCCGGGTCCTGCAGAGGTTTTCACCAAGGTCAAGG 661
Db 33 LeuTyrGluAsnGluAsnGlyHisIleArgLeuLeuGlnLysPheAspLysArgSerLys 52
QY 662 CAGTTTCAGAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTACACTCTT 721
Db 53 IlePheGluAsnLeuGlnAsnTyrArgLeuLeuGluTyrLysSerLysProHisThrLeu 72
QY 722 GTTCTTCCCAAGCAGCTGATCTGTATACATCTTGTATTATCCAGCAAGCGCAAGCCACC 781
Db 73 PheLeuProGlnTyrThrAspAlaAspPheIleLeuValValLeuSerGlyLysAlaThr 92
QY 782 GTGACGCTAGCAAAATGGCAATACAGAAAGAGCTTTAATCTTCAGCAGGCGCCATGCATC 841
Db 93 LeuThrValLeuLysSerAsnAspArgAsnSerPheAsnLeuGluArgGlyAspAlaIle 112
QY 842 AGAATCCCATCCCGTTTCATTTCTTACATCTTGAACCGCCATGACAAACAGAACCTCAGA 901
Db 113 LysLeuProAlaGlySerIleAlaTyrPheAlaAsnArgAspAsnGluGluProArg 132
QY 902 GTAGCTAAATCTCCATGCCCCGTTAAACACACCCGCGCAGTTTTCAGGATTTCTTCCCGCG 961
Db 133 ValLeuAspLeuAlaIleProValAsnLysProGlyGlnLeuGlnSerPheLeuLeuSer 152
QY 962 ACAGCGGAGAGCAATCATCTTCTTTCAGGCGCTTCAGCAGGAAATACGTTGGAGGCGGCC 1021
Db 153 GlyThrGlnAsnGlnLysSerSerLysSerGlyPheSerLysAsnIleLeuGluAlaIle 172
QY 1022 TTCAATGCGGAATTCAAATGAGATACGAGGCGTGTGTAGAAGAGAAATCAGGAGGTGAG 1081
Db 173 PheAsnThrAsnTyrGluIleGluIleGluLysValLeuLeuGluGln-----Glu 189
QY 1082 CAAGAGGAGAGGCGCAGGCGATGGAGTACTCGGAGTAGTGAG---AACATGAAGGA 1138
Db 190 GlnGluProGlnHisArgArgSerLeuLysAspArgGlnGlnIleAsnGluGluAsn 209
QY 1139 GTGATAGTCAAGTGTCAAGAGGAGCAGCTTGAAGAATCTTACTAAGCAGCCCTAAATCCGTC 1198
Db 210 ValIleValLysValSerArgGlnIleGluGluLeuSerLysAsnAlaLysSerSer 229
QY 1199 TCAGAGAAAGGCTCCGAGCAAGAGGAGGAGATATCACCAACCCCAATCAACTTGAAGAGGCG 1258
Db 230 SerLysLysSerValSerSerGlu-----SerGlyProPheAsnLeuArgSerArg 246
```



```
QY 1259 GAGCCGACTCTTCTTAACTTTGGAGCTTATTGAGGTGAAGCCAGACAAGAAGAAC 1318
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
247 AsnProIleTyrSerAsnLysPheGlyLysPheGluIleThrProGlu---LysAsn 265
QY 1319 CCCAGCTTCAGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAAGAGAGCTTGG 1378
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
266 GlnGlnLeuGlnAspLeuAspIlePheValAsnSerValAspIleLysValGlySerLeu 285
QY 1379 ATGCTCCACACTTCAACTCAAGGCGCATGGTTATGTCGTCTCAACAAGGAAGTGGGA 1438
Db ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
286 LeuLeuProAsnTyrAsnSerArgAlaIleValIleValIleValIleThrGluGlyLysGly 305
QY 1439 AACCTTGAACCTCGCTGCTGAAGAAAGACCAACAACAGAGGGGACGGCGGGAAGAGAG 1498
Db ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
306 AspPheGluLeuValGlyGlnArgAsnGluAsnGlnGlyLys-----GluAsn 321
QY 1499 GAGGACGAAGACGAAGAAGAGGAGGAGTAAACAGAGAGTGGCTAGGTACACAGCGAGG 1558
Db ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
322 AspLysGluGluGlnGluGlnGluThrSerLysGlnValGlnLeuTyrArgAlaLys 341
QY 1559 TTGAAGGAGGCGATGTGTTTCATCATGCCAGAGAGCTCATCCAGTAGCATCAACGCTTCC 1618
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
342 LeuSerProGlyAspValPheValIleProAlaGlyHisProValAlaIleAsnAlaSer 361
QY 1619 TCCGAACCTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACACACAGAACTTCTCTT 1678
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
362 SerAspLeuAsnLeuIleGlyLeuGlyIleAsnAlaGluAsnGlnArgAsnPheLeu 381
QY 1679 GCAGGTGATGAAGACATGTGATAGACCAGATAGAGAAGCAAGCAAGGATTTAGCATTC 1738
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
382 AlaGlyGluGluAspAsnValIleSerGlnValGluArgProValLysGluLeuAlaPhe 401
QY 1739 CTGGGTGCGGTGAACAAGTTGAGAAG 1765
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
402 ProGlySerSerHisGluValAspArg 410

RESULT 12
JQ2264
canavalin - jack bean
C:Species: Canavalia ensiformis (jack bean)
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 05-Oct-2004
C:Accession: JQ2264; S19137
R:Ng, J.D.; Ko, T.P.; McPherson, A.
Plant Physiol. 101, 713-728, 1993
A:Title: Cloning, expression, and crystallization of jack bean (Canavalia ensiformis) ca
A:Reference number: JQ2264; MUID:94143475; PMID:8310055
A:Accession: JQ2264
A:Molecule type: mRNA
A:Residues: 1-445 <NGJ1>
A:Cross-references: UNIPROT:P50477; UNIPARC:UPI0000110BB4; GB:S76871; NID:g17
A:Experimental source: seed
R:Ng, J.D.; Stinchcombe, T.; Ko, T.P.; Alexander, E.; McPherson, A.
Plant Mol. Biol. 18, 147-149, 1992
A:Title: PCR cloning of the full-length cDNA for the seed protein canavalin from the jac
A:Reference number: S19137; MUID:92119225; PMID:1731967
A:Accession: S19137
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-445 <NGJ2>
A:Cross-references: UNIPARC:UPI0000110BB4; EMBL:X59467; NID:g17976; PIDN:CAA42075.1; PID
C:Superfamily: vicillin
C:Keywords: storage protein

Alignment Scores:
Pred. No.: 2.63e-59 Length: 445
Score: 962.50 Matches: 198
Percent Similarity: 54.0% Conservative: 117
Best Local Similarity: 34.0% Mismatches: 101
Query Match: 26.8% Indels: 167
DB: 2 Gaps: 8

US-10-728-323-1 (1-2032) x JQ2264 (1-445)
```

```
QY 68 CCATGATGCTGTCTGTAGGATCCTTGTCTCTGGCTTCAGTTTCTGCAACGATGCCAAG 127
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
8 ProLeuTyrLeuLeuGlyValLeuLeuAlaSerValSerAlaSerPheAlaHis 27
QY 128 TCATCACCTTACCAGAAGAAAACAGAAACCCCTCGGCCAGAGGTGCTCCAGAGTTGT 187
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
28 Ser----- 28
QY 188 CAACAGAACCGGATGACTTGAAGCAAAAGGCATCGAGTCTCTGCTGCCAACCAAGTCGAG 247
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
28 ----- 28
QY 248 TATGATCCTCTGTTGTCTTATGATCCTCGAGGACACACTGTCACCAACCAACCAAGCTTCC 307
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
29 -----GlyHisSerGly----- 32
QY 308 CTTCCAGGGAGCGGACACGTCGCGCCNAACCCGGAGACTACGATGATGATCAGCGCGTCAA 367
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
33 -----GlyGluAlaGluAspGluSerGluGlu 41
QY 368 CCCCAGAGAGAGGAGGCGCGATGGGGACCGAGCTGGACCGGAGGCGTGAAGAGAA 427
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
42 SerArg----- 43
QY 428 GAAGACTGGAGACAACCAAGAGAAGATTGGAGCGCACAAAGTCAATCAGCAGCAGCCAGAAA 487
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
43 ----- 43
QY 488 ATAAGCCCCGAAGGAAGAGAAGAGACAAGAGTGGGGAACACCAGGTAGCCATGTGAGG 547
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
43 ----- 43
QY 548 GAAGAAACATCTCGGAACAACCCCTTCTACTTCCGTCAAGCGGTTTAGCACCCGCTAC 607
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
44 -----AlaGlnAsnAsnProTyrLeuPheArgSerAsnLysPheLeuThrLeuPhe 60
QY 608 GGGACACAAAACGTTAGGATCCGGGTCTCTGAGAGGTTTGACCAAAAGGTCAAGGCAGTTT 667
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 LysAsnGlnHisGlySerLeuArgLeuLeuGlnArgPheAsnGluAspThrGluLysLeu 80
QY 668 CAGAATCTCAGAATCACCGTATTGTGAGATCGAGGCCAAACCTTAACACTTGTGTTCTT 727
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
81 GluAsnLeuArgAspTyrArgValLeuGluTyrCysSerLysProAsnThrLeuLeuLeu 100
QY 728 CCCAAGCAGCTCATCTCTATAACATCTTGTATCCAGCAAGGCAAGCCACCGTGACC 787
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
101 ProHisHisSerAspSerLeuLeuValLeuGluGlyGlnAlaIleLeuVal 120
QY 788 GTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTTGACGAGGGCCATGCACCTCAGAATC 847
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 LeuValAsnProAspGlyArgAspThrTyrLysLeuAspGlnGlyAspAlaIleLysIle 140
QY 848 CCATCCGGTTTCATTTCTACATCTTGAACCCGCATGACACACAGAACCTCAGAGTAGCT 907
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
141 GlnAlaGlyThrProPheTyrLeuIleAsnProAspAsnAsnGlnAsnLeuArgIleLeu 160
QY 908 AAAATCTCCATGCCCGTTAACACACCCCGGCAGTTTGTGAGATTCTTCCCGGCGAGCAGC 967
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
161 LysPheAlaIleThrPheArgArgProGlyThrValGluAspPhePheLeuSerSerThr 180
QY 968 CGAGACCAATCATCTTACTTTGAGGGCTTCAGAGGAATACGTTGGAGGCGCGCTTCAAT 1027
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
181 LysArgLeuProSerTyrLeuSerAlaPheSerLysAsnPheLeuGluAlaSerTyrAsp 200
QY 1028 GCGGAATTCATGATACGAGGAGGTGCTGTTAGAAGAGATCCAGAGGTGAGCAGAG 1087
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
201 SerProTyrAspGluIleGluGlnThrLeuLeuGln----- 212
QY 1088 GAGAGAGGCGAGGCGATGCTACTCGGAGTAGTGAGAACAAATGAAGAGGTGATAGTC 1147
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
213 -----GluGlnGlnGluGlyValIleVal 220
QY 1148 AAAGTGTCAAAGGAGCAGCTTGAAGAACTTACTAAGCAGCGCTAAATCCGCTCAAAGAAA 1207
```


C; Keywords: seed

Alignment Scores:
 Pred. No.: 4.26e-59 Length: 445
 Score: 959.50 Matches: 197
 Percent Similarity: 54.0% Conservative: 118
 Best Local Similarity: 33.8% Mismatches: 101
 Query Match: 26.8% Indels: 167
 DB: 2 Gaps: 8

US-10-728-323-1 (1-2032) x S00281 (1-445)

Qy	68	CCACTGATGCTGTGCTAGGATCCTTGTCTCTGGCTTCAGTTTCTGCAACGATGCCAAG	127
Db	8	ProLeuTrpLeuLeuGlyValValLeuLeuAlaSerValSerAlaSerPheAlaHis	27
Qy	128	TCATCACCTTACCAGAGAAACACAGAACCCCTCGCGCCAGAGGTGCCTCCAGAGTTGT	187
Db	28	Ser-----	28
Qy	188	CAACAGAACCGGATGACTTTGAAGCAAAAGGCATGCGAGTCTCGCTGCACCAAGCTCGAG	247
Db	28	-----	28
Qy	248	TATGATCCTCGTTGTGTATGATCTCTGAGGACACACTGGCACCCACCAACCAACGTTCC	307
Db	29	-----GlyHisSerGly-----	32
Qy	308	CCTCCAGGGAGCGGACACGTTGGCCGCCAACCCGGAGACTACGATGATGACCGCGTCNA	367
Db	33	-----GlyGluAlaGluAspGluSerGluGlu	41
Qy	368	CCCGAAGAGAGGAAGGAGGCGATGGGACCAAGCTGGACCGGAGGCGGTGAAGAGAA	427
Db	42	SerArg-----	43
Qy	428	GAAGACTGGAGACAACCAAGAGAAGATTGGAGGCGACAAGTCACTACGACGCCGGA	487
Db	43	-----	43
Qy	488	ATAAGGCCCAAGGAAGAGAGGAGACAAGAGTGGGGAACACACAGGTAGCCATGTGAGG	547
Db	43	-----	43
Qy	548	GAAGAAACATCTCGAAACAACCTTTCTACTTCCCGTCAAGGCGGTTTAGCACCCGCTAC	607
Db	44	-----AlaGlnAsnAsnProTyrLeuPheArgSerAsnLysPheLeuThrLeuPhe	60
Qy	608	GGGAACCAAAACGGTAGGATCCGGGTCCTCGAGAGTTTGACCAAGGTCAAGGCAGTTT	667
Db	61	LysAsnGlnHisGlySerLeuArgLeuGlnArgPheAsnGluAspThrGluLysLeu	80
Qy	668	CAGAACTTCAGAAATCACCGTATTGTGTCAGATCGAGGCCAACCTTAACACTCTTCTT	727
Db	81	GluAsnLeuArgAspTyrArgValLeuGluTyrCysSerLysProAsnThrLeuLeu	100
Qy	728	CCAAGCACCTGATGCTGTATAACATCCTTGTATTCCAGCAAGGCAACCCACCGTGACC	787
Db	101	ProHisSerAspSerAspLeuValLeuValLeuGluGlyGlnAlaLeuVal	120
Qy	788	GTACCAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAGGCGCATGCATCAGAATC	847
Db	121	LeuValAsnProAspGlyArgAspThrTyrLysLeuAspGlnGlyAspAlaIleLysIle	140
Qy	848	CCATCCGGTTTCATTTCCCTACATCTTGAAACCGCCATGACACACAGAACCTCAGATAGCT	907
Db	141	GlnAlaGlyThrProPheTyrLeuIleAsnProAspAsnAsnGlnAsnLeuArgIleLeu	160
Qy	908	AAAATCCCATGCCGTTAAACACACCCGGCCAGTTTGAGGATTCTTCCCGCGCAGCAGC	967
Db	161	AsnPheAlaIleThrPheArgArgProGlyThrValGluAspPhePheLeuSerThr	180
Qy	968	CGAGACCAATCATCTACTTTCGCGGGCTTCAGCAGGAATACGTTGGAGCGCGCTTCAAT	1012

Db	181	Ly	Ser	Arg	Leu	Pro	Ser	Tyr	Leu	Ser	Ala	Phe	Ser	Lys	Asn	Phe	Leu	Ala	Ser	Tyr	Asp	200	
		:::																				200	
Qy	1028	CG	GAA	TTC	CAAT	GAG	AT	CAC	GAG	GGT	GCT	GTT	TAG	AAG	AAG	AAT	GAG	GAG	GTT	GAG	CA	1087	
																					1087		
Db	201	Ser	Pro	Tyr	Arg	Asp	Gl	u	Ile	Leu	Leu	Gln	---	---	---	---	---	---	---	---	---	212	
																					212		
Qy	1088	GAG	AG	AGG	G	AGG	CGG	AT	GAG	TAC	TCT	CGG	AGT	AGT	AGT	AGT	AGT	AGT	AGT	AGT	AGT	1147	
																					1147		
Db	213	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	220	
																					220		
Qy	1148	AA	AGT	GC	TCA	AA	GAG	GAC	CGT	TGA	AA	GCT	TAC	TAA	GAC	GCT	TAA	TCC	GCT	CA	AA	1207	
																					1207		
Db	221	Lys	Met	Pro	Lys	Asp	Gln	Ile	Leu	Gln	Ile	Ser	Lys	His	Ala	Gln	Ser	Ser	Ser	Arg	Lys	240	
																					240		
Qy	1208	GG	CT	CC	GA	AG	AA	GAG	GG	AG	TAT	CAC	CA	ACC	CA	TCA	ACT	TG	G	AG	AG	1267	
																					1267		
Db	241	Thr	Leu	Ser	Ser	Gln	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	257	
																					257		
Qy	1268	CT	TTC	TAA	CA	ACT	TTT	GGG	AG	AT	TAT	TTC	GAG	TGA	AGC	CAG	CA	CA	AG	AA	CA	1327	
																					1327		
Db	258	Tyr	Ser	Asn	Asn	Tyr	Gi	Lys	Leu	Tyr	Gl	Ile	Thr	Pro	---	---	---	---	---	---	---	276	
																					276		
Qy	1328	CAG	AC	CT	GG	CA	TG	AT	GC	TAC	CT	GT	G	TAG	AG	AT	CA	AA	GAG	GAG	CT	1387	
																					1387		
Db	277	Arg	Asp	Leu	Asp	Ile	Leu	Leu	Asn	Cys	Leu	Gln	Met	Asn	Gl	u	Gl	u	Ala	Leu	Phe	Val	296
																					296		
Qy	1388	CAC	TT	CA	ACT	CA	AA	G	GC	AT	GTT	TAT	CGT	CGT	CT	CA	CA	AA	G	AA	CA	1447	

```

RESULT 14
T44430
protein PV100 [imported] - winter squash
C:Species: Cucurbita maxima (winter squash)
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T44430
R:Yamada, K.; Shimada, T.; Kondo, M.; Nishimura, M.; Hara-Nishimura, I.
J. Biol. Chem. 274, 2563-2570, 1999
A:Title: Multiple functional proteins are produced by cleaving Asn-Gln bonds of a single
A:Reference number: 222767; MUID:99107919; PMID:9891029
A:Accession: T44430
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-810 <YAM>

```

A;Cross-references: UNIPROT:Q9ZWI3; UNIPARC:UPI00000A1E11; EMBL:AB019195; NID:G3808061

Alignment Scores:
Pred. No.: 2,07e-55 Length: 810
Score: 907.50 Matches: 214
Percent Similarity: 54.6% Conservative: 121
Best Local Similarity: 34.9% Mismatches: 191
Query Match: 25.3% Indels: 87
DB: 2 Gaps: 20

US-10-728-323-1 (1-2032) x T44430 (1-810)

QY 188 CAACAGGAACCGGAT--GACTTCAAGCAAAAGGCATCGGAGTCTCGTGCACCAAGCTC 244
DB 244 GlnArgAspProAspTrpArgGluGlnIysArgArgGluGlnGluGlnArgArg 263
QY 245 GAGTATGATCTCTGGTGTCTATGATCTCTCGAGACACACTGGCACCACCAACGTCGT 304
DB 264 GluArgGluArgGlyGlyArgAspGluAspGlu-----AsnGlnArg 279
QY 305 TCCCTCCAGGGAGCGACACGTGGCGCCCAACCCGGAGAC--TACGATGATCACGCG 361
DB 280 AspProAspTrpArgGluGlnGluArgGluGlnGluArgArgArgGluArg 299
QY 362 CGTCAACCCCGAAGAGAGAGAGCGGATGGGACCGAGCTGGACCGAGGAGCGTGA 421
DB 300 GluGlnGluArgGluArgGluHisArgGlyArgAspValGluAspGluAsnGln 319
QY 422 AGAAGAGAGACTGGAGACACCAAGAGAAATGGAGCGACCAAGTCATCAGCAGCCA 481
DB 320 ArgAspProAspTrpArgGluGlnGluArgArgGluGluGlnArgGlu 339
QY 482 CGGAAA---ATAAGGCCCGAAGAGAGAGAGAGAAAGAGTGGGAAACACCAAGTAGC 538
DB 340 ArgGluTrpGluArgGluHisGlyArgGlyArgGluGlnArgSerArgGluAsp 359
QY 539 CATGTGAGG----- 547
DB 360 GluArgArgArgHisGluArgGlnHisGlyArgSerArgValAsnGlnValAlaIle 379
QY 548 -----GAAGAAACATCTCGGAACAACCTTCTACTTCCCGTCAAGGCGGTTAGC 598
DB 380 ArgArgThrGluGlnGluGlnSerAsnAsnProTyrPheGlnGluGlnArgPheGln 399
QY 599 ACCCGCTACGGAAACCAAAACGGTAGGATCCGGTCTCGAGAGTTTGACCAAGGTCA 658
DB 400 SerArgTyrArgSerArgGluGlyHisTrpValLeuGluArgPheSerGluArgSer 419
QY 659 AGGCAGTTTCAGATCTCCGAATCACCGTATTGTGCAGATCGAGGCCAAACCTTAACCT 718
DB 420 GluLeuLeuIysGlyIleIysAsnGlnArgIleuAlaLeuLeuGluAlaArgProHisThr 439
QY 719 CTGTGTTCTTCCCAAGCAGCTGATGCTGATAACATCTTGTATTATCAGCAGGGCAAGCC 778
DB 440 PheIleValProHisHisLeuAspAlaGluCysValLeuLeuValValArgGlyArgAla 459
QY 779 ACCGTGACCGTAGCA-----AATGCCAATAACAGAAAGAGCTTTAATCTTGAC 826
DB 460 ThrIleThrValValIleGlnIleIysArgGluThrArgLysGluSerTyrAsnValGlu 479
QY 827 GAGGCCCATGCACTCAGAAATCCATCCGGTGTTCATTTCTCATCTTCAACCGCCATGAC 886
DB 480 SerGlyAspValMetThrIleProIleGlyThrThrLeuTyrLeuAlaAsnGln---Glu 498
QY 887 AACCAAGAACCTCAGAGTAGCTAAATCTCCATGCCCCGTAAACACACCCGCGCAGTTTGAG 946
DB 499 AsnGluAspLeuGlnIleValIysLeuValGlnProValAsnAsnProGlyIlePheLys 518
QY 947 GATTCTCTCCCGCGAGCAGCCGAGACCAATCATCTTCTGACGGCTTTCAGCAGGAAT 1006
DB 519 AspTyrLeuSerAlaGlyGlyGluSerGln---AlaTyrTyrSerValPheSerAsnAsp 537
QY 1007 ACGTTGGAGCGCCCTCAATGCGGAATTCATATGATACATCGGAGGGTGTGTAGAAAGAG 1066

Db 581 AspArgArgGlyPheAspGluArgArgGlySerAsnAsnProLeuSerPro 597

Search completed: May 15, 2006, 22:23:08
Job time : 109.003 secs

Db	241	GlulysileTyrValValThrAsnGlyArgGlyThrValThrPheValThrHisGluAsn	250
Qy	806	AGAAAGAGCTTTAATCTTGACGAGGGCCATGCACCTCAGATCCCATCCGGTTCATTCTCC	865
Db	261	LysGluSerTyrAsnValProGlyValValArgIleProAlaGlySerThrVal	280
Qy	866	TACATCTTGAACCGCCATGACACCAACCACTCAGTAGCTAAAATCTCCATGCCGGTT	925
Db	281	TyrLeuAlaAsnGlnAspAsnArgGluLysLeuThrIleAlaValLeuHisArgProVal	300
Qy	926	AACACACCCGGCCAGTTTGAGGATTTCTTCCCGCGAGCAGCCGACCAATCATCTCTAC	985
Db	301	AsnAsnProGlyGlnPheGlnLysPhePheProAlaGlyGlnGluAsnProGlnSerTyr	320
Qy	986	TTGCAGGGCTTCACGACGAATACGTTGAGAGCCCTTCAATCGGNAATTCATGAGATA	1045
Db	321	LeuArgIlePheSerArgGluIleLeuGluAlaValPheAsnThr	335
Qy	1046	CGGAGGGTGTGTTAGAAAGAGATGCAAGAGAGGTGAGCAAGAGAGAGAGCGCAGAGCGGA	1105
Db	336	ArgSerGluGlnLeuAspGluLeuProGlyGlyArgGlnSerHisArgArgGlnGln---	354
Qy	1106	TGGAGTACTCGAGTAGTGAGAACAAATGAAGGAGTAGTATCAAGTGTCAAAGGAGCAC	1165
Db	355	-----GlyGlnGlyMetPheArgIysAlaSerGlnGluGln	366
Qy	1166	GTTGAAGNACTTACTAAGCAGCGCTAAATCCGTCTCAAGAAAGCGTCCGAGAGAGGGA	1225
Db	367	IleArgAlaLeuSerGlnGlyAlaThrSerProArgGlyLysGlySerGluGlyTyrAla	386
Qy	1226	GATATCACCAACCCAACTCAACTTCAGAGAAGCGAGCCCGATCTTCTAAACAACCTTGGG	1285
Db	387	-----PheAsnLeuLeuSerGlnThrProArgTyrSerAsnGlnAsnGly	401
Qy	1286	AAGTTATTTGAGTGAAGCCAGCAAGAAACCCCGAGTTCAGGACCTTGGACATGATG	1345
Db	402	ArgPheTyrGluAlaCysProArgAsnPheGlnGlnGlnLeuArgGluValAspSerSer	421
Qy	1346	CTCACCTGTGTAGAGATCAAGAAGGAGCTTTGATGTCTCCACACTTCACTCAAGGCC	1405
Db	422	ValValAlaPheGluIleAsnLysGlySerIlePheValProHisTyrAsnSerLysAla	441
Qy	1406	ATGGTTATCGTCGTCGTCACAAAGGAACTTGAACCTTCAACTCGTGGCT-----GTA	1459
Db	442	ThrPheValValLeuValThrGluGlyAsnGlyHisValGluMetValCysProHisLeu	461
Qy	1460	AGAAAAGAGCAACACAGAGGGGACGGGAGAGAGAGGAGGACGAAGAAGAGAG	1519
Db	462	SerArgGlnSerSerAspTyrSerArgGluGluGlnGluGluGlnGluGlnVal	481
Qy	1520	GAGGGAAGTACAGAGAGGTCGTAGGTACACAGAGGTTGAAGGAAGCGGATGTCTTC	1579
Db	482	GluArgArgSerGlyGlnTyrLysArgValArgAlaGlnLeuSerThrGlyAsnLeuPhe	501
Qy	1580	ATCATGCCAGAGCTCATCCAGTAGCCATCAACCGCTTCC-----TCGGAACCTCATCTG	1633
Db	502	ValValProAlaGlyHisProValThrPheValAlaSerGlnAsnGluAspLeuGlyLeu	521
Qy	1634	CTTGGCTTCGGTATC---AAGCGTTGAAAACCAACCAAGATCTTCTTTCGAGGTGATAAG	1690
Db	522	LeuGlyPheGlyLeuTyrAsnGlyGlnAspAsnLysArgIlePheValAlaGly---Lys	540
Qy	1691	GACAAATGTGATAGACCATAGAGAACGACGAGGATTTAGCATTTCCCTGGGTCCGGT	1750
Db	541	ThrAsnAsnValArgGlnTrpAspArgGlnAlaLysGluLeuAlaPheGlyValGluSer	560
Qy	1751	GAAACAAGTTGAGAAGCTCATCAAAAACCAAG---AAGGAATCTCACTTTGTGAGTCTCGT	1807
Db	561	ArgLeuValAspGluValPheAsnAsnAsnProGlnGluSerTyrPheValSerGlyArg	580
Qy	1808	-----CCTCAATCTCAATCTCAATCTCCGTCGTCTCTCT	1840

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 21:23:02 ; Search time 21.4781 Seconds
(without alignments)
4400.310 Million cell updates/sec

Title: US-10-728-323-2

Perfect score: 1281

Sequence: 1 gctcaccatactagtagcccc.....taaaagatcatgtttgtt 717

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p model -DEV=xlp
-Q-/abs/ABSSWEB_spool/US10728323/runat_15052006_172130_22366/app_query.fasta.1
-DB=A_Geneseq -Qfmt=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=US10728323 @CCN 1 1 605 @runat 15052006 172130_22366 -NCPU=6 -ICPU=3
-NO_MAP -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: Geneseq21.*
- 2: Geneseq1980s.*
- 3: Geneseq1990s.*
- 4: Geneseq2000s.*
- 5: Geneseq2001s.*
- 6: Geneseq2002s.*
- 7: Geneseq2003as.*
- 8: Geneseq2003bs.*
- 9: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	842	65.7	157	2	Aaw24153 Peanut al
2	842	65.7	157	2	Aaw24164 Peanut al
3	842	65.7	157	2	Aay15245 Peanut al
4	842	65.7	157	4	Aau04707 Anaphylac
5	842	65.7	157	6	Abu52464 Peanut Ar
6	842	65.7	157	7	Adg27518 Peanutur al
7	842	65.7	157	9	Adv97542 Peanut Ar
8	842	65.7	159	9	Adv97603 Peanut Ar
9	838.5	65.5	207	4	Aab82383 Peanut al

10	830	64.8	156	2	AAY40973	Aay40973 Ara h 2 p
11	827	64.6	157	9	ADV97553	Peanut Ar
12	827	64.6	156	8	ADO38314	Ara h 2 M
13	814	63.9	156	6	ABU52576	Peanut Ar
14	812	63.4	156	8	ADM12096	Arachis h
15	812	63.4	157	6	ABU52575	Peanut Ar
16	774	60.4	157	6	ABU52577	Peanut Ar
17	771	60.2	166	2	AAY40968	Aay40968 Recombina
18	771	60.2	166	6	ABU52482	Peanut Ar
19	771	60.2	166	7	ADG27536	T7/His-ta
20	699	54.6	166	3	AA333600	Modified
21	699	54.6	166	3	AA333600	Modified
22	688.5	53.7	167	4	AAU04710	Modified
23	243	19.0	158	2	AAW23419	Soybean a
24	243	19.0	158	2	AAW23419	Soybean G
25	243	19.0	158	7	ADH89269	G. max 2S
26	243	19.0	158	8	ADG44004	G. max 2S
27	241.5	18.9	155	2	AAW23418	Soybean a
28	241.5	18.9	155	7	ADH89267	G. max na
29	241.5	18.9	155	8	ADG44002	G. max 2S
30	229	17.9	158	2	AAW23420	Chimeric
31	184.5	14.4	323	7	ADH89277	Sunflower
32	184.5	14.4	323	8	ADG44012	H. annuus
33	177.5	13.9	168	6	AAE36076	Flax Conl
34	156	12.2	28	2	AAW24194	Peanut al
35	156	12.2	28	9	ADV97585	N-termina
36	156	12.2	158	2	AAW23586	Mablinin
37	151	11.8	158	2	AAW23588	Mablinin
38	151	11.8	295	7	ADH89275	Sunflower
39	151	11.8	295	8	ADG44010	H. annuus
40	149.5	11.7	167	8	ADS16393	Grape 2S
41	146.5	11.4	164	1	AAW23586	1kb fragm
42	146.5	11.4	164	1	AAW23586	1kb fragm
43	146.5	11.4	164	1	AAW23586	1kb fragm
44	146.5	11.4	164	1	AAW23586	1kb fragm
45	144	11.2	169	6	AAE36077	Flax Conl

ALIGNMENTS

RESULT 1

AAW24153
ID AAW24153 standard; protein; 157 AA.

- XX AC AAW24153;
- XX AC AAW24153;
- DT 17-OCT-2003 (revised)
- DT 29-DEC-1997 (first entry)
- XX
- DE Peanut allergen Ara hII.
- XX
- XX Peanut; seed storage protein; allergen; hypersensitivity;
- KW vaccine; anaphylactic shock; immunotherapy; therapy; monoclonal antibody;
- KW ELISA; analysis; Ara hII.
- XX

OS Arachis hypogaea; strain Florunner.

PN WO9724139-A1.

PD 10-JUL-1997.

PF 23-SEP-1996; 96WO-US015222.

PR 29-DEC-1995; 95US-0009455P.

PR 04-MAR-1996; 96US-00610424.

(UYAR-) UNIV ARKANSAS.

Burks AW, Helm RM, Cockrell G, Stanley JS, Bannon GA;

WPI; 1997-363453/33.

DR N-PSDB; AAT76615.

XX

PT Peanut allergens Ara hi and Ara hII - used for vaccination and in two-site monoclonal antibody based ELISA.
 XX
 XX Claim 31; Page 198; 354pp; English.
 XX
 CC This polypeptide comprises major peanut allergen Ara hII. Its sequence was deduced from a cDNA clone (AA76615) isolated from peanut seed cDNA using a primer (see AAT76617) based on an isolated Ara hi peptide (see AAW24151). The sequence shows significant homology with the conglutin family of seed storage proteins of other legumes. The allergen is recognised by serum IgE from a large proportion of individuals with peanut hypersensitivity. Ara hII and Ara hi (see AAW24149-50) can be used to raise monoclonal antibodies which are used in a specific two-site MAB ELISA for the detection of Ara hi or Ara hII (claimed). IgE-binding Ara hII antigen epitopes (see AAW24188-93) may be used in vaccines to protect against allergic reactions to peanut allergens, e.g. anaphylactic shock.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 XX
 XX Sequence 157 AA;

Alignment Scores:
 Pred. No.: 9,32e-90 Length: 157
 Score: 842.00 Matches: 157
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 65.7% Indels: 0
 DB: 2 Gaps: 0

US-10-728-323-2 (1-717) x AAW24153 (1-157)

QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCGCTCGAGGCGAG 61
 DB 1 LeuThrIleuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
 QY 62 CAGTGGGAACCTCAAAGGAGACAGAGATGCCAGACCGCTCGAGAGGGGCAACCTGAGG 121
 DB 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
 QY 122 CCTGTGCGAACAATCTTCATGCAGAGATCCAAAGTGCAGGAGTTCATATGAACGGGAC 181
 DB 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
 QY 182 CCGTACAGCCCTAGTAGCGCTAGCGCCCTAGTCCATATGTCGAGAGCGCTGGA 241
 DB 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
 QY 242 TCCTCTCAGCACCAAGAGAGGTGTTCATGATGAGTGAACGAGTTTGAGAACCAACCAAGG 301
 DB 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100
 QY 302 TGCATGTGCGAGGCAATTCGACAGATCATGCGAGAACCGAGCGGATAGTTGCGAGGGAGG 361
 DB 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
 QY 362 CACAGAGCAACAGTTCAAGAGGAGCTCAGGAACTTGCTCAACAGTCGCGCTTAGG 421
 DB 121 GlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
 QY 422 GCACCACAGCGTTGCGACTTGGAGCTCGAAAGTGGCGGCGAGACAGATAC 472
 DB 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157

RESULT 2

AAW24164
 ID AAW24164 standard; protein; 157 AA.

XX AC AAW24164;

XX DT 17-OCT-2003 (revised)

XX DT 29-DEC-1997 (first entry)

XX XX Peanut allergen Ara hII.

DE
 XX

KW Peanut; seed storage protein; allergen; allergy; hypersensitivity;
 KW vaccine; anaphylactic shock; immunotherapy; therapy; monoclonal antibody;
 KW ELISA; analysis; Ara hII.

XX Arachis hypogaea; strain Florunner.

XX WO9724139-A1.

XX 10-JUL-1997.

XX 23-SEP-1996; 96WO-US015222.

XX 29-DEC-1995; 95US-0009455P.

XX 04-MAR-1996; 96US-00610424.

XX (UYAR-) UNIV ARKANSAS.

XX Burks AW, Helm RM, Cockrell G, Stanley JS, Bannon GA;

XX WPI; 1997-363453/33.

XX N-PSDB; AAT76614.

XX Peanut allergens Ara hi and Ara hII - used for vaccination and in two-site monoclonal antibody based ELISA.

XX Claim 31; Page 219; 354pp; English.

CC This polypeptide comprises major peanut allergen Ara hII. Its sequence was deduced from cDNA clone P38 (AAT76614), isolated from peanut seed cDNA using a primer (see AAT76617) based on an isolated Ara hi peptide (see AAW24151). The sequence shows significant homology with the conglutin family of seed storage proteins of other legumes. The allergen is recognised by serum IgE from a large proportion of individuals with peanut hypersensitivity. Ara hII and Ara hi (see AAW24149-50) can be used to raise monoclonal antibodies which are used in a specific two-site MAB ELISA for the detection of Ara hi or Ara hII (claimed). IgE-binding Ara hII antigen epitopes (see AAW24188-93) may be used in vaccines to protect against allergic reactions to peanut allergens, e.g. anaphylactic shock.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 XX

XX Sequence 157 AA;

Alignment Scores:
 Pred. No.: 9,32e-90 Length: 157
 Score: 842.00 Matches: 157
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 65.7% Indels: 0
 DB: 2 Gaps: 0

US-10-728-323-2 (1-717) x AAW24164 (1-157)

QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCGCTCGAGGCGAG 61
 DB 1 LeuThrIleuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
 QY 62 CAGTGGGAACCTCAAAGGAGACAGAGATGCCAGACCGCTCGAGAGGGGCGACCTGAGG 121
 DB 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
 QY 122 CCTGTGCGAACAATCTTCATGCAGAGATCCAAAGTGCAGGAGTTCATATGAACGGGAC 181
 DB 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
 QY 182 CCGTACAGCCCTAGTAGCGCTAGCGCCCTAGTCCATATGTCGAGAGCGCTGGA 241
 DB 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
 QY 242 TCCTCTCAGCACCAAGAGAGGTGTTCATGATGAGTGAACGAGTTTGAGAACCAACCAAGG 301
 DB 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100
 QY 302 TGCATGTGCGAGGCAATTCGACAGATCATGCGAGAACCGAGCGGATAGTTGCGAGGGAGG 361

Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAasnGlnSerAspArgLeuGlnGlyArg 120
 QY 362 CAACAGGAGCAACAGTTCAGAGGAGCTCAGGAACCTGCCTCAACAGTGGCGCCTTAGG 421
 Db 121 GlnGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
 QY 422 GCACACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGAGAGACACAGATAC 472
 Db 141 AlaProGlnArgCysAspLeuaspValGluSerGlyGlyArgAspArgTyr 157
 RESULT 3
 AAY15245
 ID AAY15245 standard; protein; 157 AA.
 XX
 AC AAY15245;
 XX
 DT 17-OCT-2003 (revised)
 DT 09-NOV-1999 (first entry)
 XX
 DE Peanut allergen, Ara h 2, amino acid sequence.
 XX
 DE
 KW allergy; immune response; transgenic; allergen; epitope;
 KW immunoglobulin E; Ig E; binding site; peanut.
 XX
 OS Arachis hypogaea.
 XX
 PN WO9938978-A1.
 XX
 PD 05-AUG-1999.
 XX
 PF 29-JAN-1999; 99WO-US002031.
 XX
 PR 31-JAN-1998; 98US-0073283P.
 PR 13-FEB-1998; 98US-0074590P.
 PR 13-FEB-1998; 98US-0074624P.
 PR 13-FEB-1998; 98US-0074633P.
 PR 27-AUG-1998; 98US-00141220.
 XX
 PA (UYAR-) UNIV ARKANSAS.
 PA (UYNY) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.
 PA (SOSI/) SOSIN H.
 XX
 PI Sosin H, Bannan GA, Burks AW, Sampson HA;
 XX
 DR WPI: 1999-479189/40.
 DR N-PSDB; AAZ06383.
 XX
 PT Modified allergen with reduced IgE binding, useful for treating e.g.
 PT allergies.
 XX
 PS Disclosure; Page 38; 46pp; English.
 XX
 CC This is the amino acid sequence of the Ara h 2 protein from Arachis
 CC hypogaea. The Ara h 2 protein has 10 IgE (immunoglobulin E) binding
 CC epitopes, three of which are immunodominant (AAY15272, AAY15275, and
 CC AAY15276). By modifying the IgE binding sites the ability of the allergen
 CC to provoke an immune response is downregulated. The epitopes of the IgE
 CC binding sites can therefore be modified in genetically engineered plants
 CC and animals to elicit less of an allergic response. (Updated on 17-OCT-
 CC 2003 to standardise OS field)
 XX
 SQ Sequence 157 AA;
 Alignment Scores:
 Pred. No.: 9.32e-90 Length: 157
 Score: 842.00 Matches: 157
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 65.7% Indels: 0
 DB: 2 Gaps: 0

US-10-728-323-2 (1-717) x AAY15245 (1-157)

QY 2 CTCACATAGTAGCCCTCGCCCTTTTCTCTCTCGCTGCCACGCATCTCGAGGCAG 61
 Db 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaIleAlaSerAlaArgGln 20
 QY 62 CAGTGGGAACCTCAAGGAGACAGAAGATGCCAGACCCAGCTCGAGAGCGCGAACCTGAGG 121
 Db 21 GlnTTPGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
 QY 122 CCCTCGGAGCACATCTCATGCAGAGATCCCAACCTGACGAGGATTTCATATGAACCGGAC 181
 Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgaspGluaspSerTyrGluArgasp 60
 QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 241
 Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
 QY 242 TCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAACCAACAAAGG 301
 Db 81 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnGlnArg 100
 QY 302 TGCATGTGGAGGCATTGCACAGATCATGGAGAACCCAGAGCGATAGTTGCGAGGGAGG 361
 Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
 QY 362 CAACAGGAGCAACAGTTCAGAGAGGAGCTCAGGAACTTGCCTCAACAGTCCGCGCTTAGG 421
 Db 121 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
 QY 422 GCACACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGAGAGACAGATAC 472
 Db 141 AlaProGlnArgCysAspLeuaspValGluSerGlyGlyArgaspArgTyr 157

RESULT 4

AAU04707

ID AAU04707 standard; protein; 157 AA.

XX AAU04707;

DT 23-OCT-2001 (first entry)

XX Anaphylactic antigen Ara h 2.

KW Ara h 2; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;

KW allergy; mast cell; basophil; mouse.

XX Mus sp.

XX WO200140264-A2.

XX 07-JUN-2001.

XX 06-DEC-2000; 2000WO-US033124.

XX 06-DEC-1999; 99US-00455294.

PR 23-JUN-2000; 2000US-0213765P.

PR 27-SEP-2000; 2000US-0235797P.

XX (PANA-) PANACEA PHARM LLC.

PA (UYAR-) UNIV ARKANSAS.

PA (MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.

XX

PI Bannan GA, Burks WA, Caplan MJ, Sampson H, Sosin H;

XX WPI; 2001-381378/40.

DR

XX

PT Antigenic fragments useful for reducing anaphylactic risk and reducing

PT the severity and/or number of allergic symptoms in individuals sensitive

XX to antigens, have reduced ability to bind Immunoglobulin E.

XX Claim 7; Fig 10; 100pp; English.

XX

CC The sequence represents the amino acid sequence of anaphylactic antigen

CC Ara h 2. Ara h 2 is an anaphylactic antigen (A), which was used to design
CC antigenic peptides having a reduced ability to bind immunoglobulin E
CC (IgE) as compared with the intact (A), or having a sequence substantially
CC identical to a portion of sequence of an antigen that includes at least
CC one IgE binding site, where at least one IgE binding site of the peptide
CC is altered. The antigenic peptides are used in a composition which is
CC useful for reducing risk or severity of allergic reaction to an antigen.
CC This is done by identifying an individual at risk of allergic reaction to
CC an antigen by identifying prior display of allergic symptoms when exposed
CC to the antigen, or a familial relationship with an individual who
CC previously displayed allergic symptoms when exposed to the antigen.
CC Following this an antigen-specific IgE present on one or more mast cells
CC or basophils in the individual's serum is identified. The individual is
CC then contacted with a peptide corresponding to a portion of the antigen,
CC which is selected, formulated, and delivered so that binding of the
CC peptide to antigen-specific IgE is reduced as compared with IgE binding
CC of intact antigen. The composition is also useful for treating and
CC preventing allergic reactions
XX
SQ Sequence 157 AA;

Alignment Scores:
Pred. No.: 9.32e-90 Length: 157
Score: 842.00 Matches: 157
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 65.7% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-2 (1-717) x AAU04707 (1-157)

QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTCGTCCACGGATCTGGAGGCAG 61
DB 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAACCTCAAGGAGACAGAGATGCCAGCCAGCTCCAGAGGGCGAACCTGAGG 121
DB 21 GlnTrpGluLeuGlnGlyAspArgGlySerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCCTGCGAGCAACATCTCATGCGAAGATCCACGTGACGAGGATTCATATGAACGGAC 181
DB 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CCGTACAGCCCTAGTACGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 241
DB 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
QY 242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAACCAAGG 301
DB 81 SerSerGlnHisGlnGluArgGlyCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100
QY 302 TGCATGTGCGAGGATTTGCAACAGATCATGGAGAACCGAGCGCATAGTTGCGAGGGAGG 361
DB 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGGAGCAAGTTCAGAGGGAGCTCAGGAACTTGCCCTCAACAGTGGCGCTTAGG 421
DB 121 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACCAGAGCGTTGCGAGCTTGAGAGCTCGAAAGTGGCGGCGAGACAGATAC 472
DB 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157

RESULT 5

ABU52464

ID ABU52464 standard; protein; 157 AA.

XX AC ABU52464;

XX 10-MAR-2003 (first entry)

XX Peanut Ara h2 protein.

XX

KW Peanut; allergy; Ara h1; Ara h2; Ara h3; IgE binding site;
KW anaphylactic food allergen; antiallergenic; vaccine; wound healing.
XX
OS Arachis hypogaea.
XX
PN WO200274250-A2.
XX
PD 26-SEP-2002.
XX
XX 18-MAR-2002; 2002WO-US009108.
XX
PR 16-MAR-2001; 2001US-0276822P.
PR 19-MAR-2002; 2002US-00276822.
XX
PA (PANA-) PANACEA PHARM.
XX
XX Caplan M, Sosin H, Sampson H, Bannan GA, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabin PA, Shin DS, Stanley JS;
XX
XX WPI; 2003-018765/01.
DR N-PSDB; ABX70606.
XX
XX New modified anaphylactic food allergen, useful for preventing or
PT treating allergic reactions associated with e.g. anaphylactic allergens.
XX
XX Claim 27; Fig 41; 300pp; English.
XX
XX The invention relates to a modified anaphylactic food allergen has an
CC amino acid sequence that is substantially identical to that of natural
CC anaphylactic food allergen, except for a cysteine residue that has been
CC modified so that it cannot participate in the disulphide bond. The
CC modification may also comprise mutation of the IgE binding sites to
CC reduce allergenicity. Also included are: (1) a nucleotide molecule encoding
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
CC or for causing a site specific mutation in the modified anaphylactic food
CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified
CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a
CC peanut allergen (e.g. Ara h1, h2 or h3)
XX
SQ Sequence 157 AA;

Alignment Scores:

Pred. No.: 9.32e-90 Length: 157
Score: 842.00 Matches: 157
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 65.7% Indels: 0
DB: 6 Gaps: 0

US-10-728-323-2 (1-717) x ABU52464 (1-157)

QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTCGTCCACGGATCTGGAGGCAG 61
DB 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAACCTCAAGGAGACAGAGATGCCAGCCAGCTCCAGAGGGCGAACCTGAGG 121
DB 21 GlnTrpGluLeuGlnGlyAspArgGlySerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCCTGCGAGCAACATCTCATGCGAGAGATCCACGTGACGAGGATTCATATGAACGGAC 181
DB 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CCGTACAGCCCTAGTACGAGTCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 241

Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
QY 242 TCCTCTCAGCACCAAGAGGTGTGCAATGAGCTGAACGAGTTTCAGAACACCAAGG 301
Db 81 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100
QY 302 TGCATGTGCGAGGCATTGCAACAGATCATGGAGAACCAAGAGCGATAGTTGCAGGGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGAACTTGCCTCAACAGTGGCGCTTAGG 421
Db 121 GlnGlnGlnGlnGlnPheIlysrGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACCCAGCGTTTGCAGCTTGGACGTGCGAAGTGGCGGAGAGACAGATAC 472
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157
RESULT 6
ADG27518
ID ADG27518 standard; protein; 157 AA.
XX
AC ADG27518;
XX
DT 26-FEB-2004 (first entry)
XX
DE Peanut allergen Ara h2.
XX
KW Peanut; plant; allergen; Ara h1; Ara h2; Ara h3; glycinin A2B1a; Jug n1;
KW antiallergic; vulnerability; anaphylactic food allergen; IgE; allergy; wound.
XX
OS Arachis hypogaea.
XX
PN US2003202980-A1.
XX
PD 30-OCT-2003.
XX
PF 18-MAR-2002; 2002US-00100303.
XX
PR 29-DEC-1995; 95US-0009455P.
PR 23-SEP-1996; 96US-00717933.
PR 31-JAN-1998; 98US-0073283P.
PR 13-FEB-1998; 98US-0074590P.
PR 13-FEB-1998; 98US-0074624P.
PR 13-FEB-1998; 98US-0074633P.
PR 29-JUN-1998; 98US-00106872.
PR 27-AUG-1998; 98US-00141220.
PR 13-NOV-1998; 98US-00191593.
PR 29-JAN-1999; 99US-00240557.
PR 29-JAN-1999; 99US-00241101.
PR 11-FEB-1999; 99US-00248673.
PR 11-FEB-1999; 99US-00248674.
PR 02-MAR-1999; 99US-0122450P.
PR 02-MAR-1999; 99US-0122452P.
PR 02-MAR-1999; 99US-0122560P.
PR 02-MAR-1999; 99US-0122565P.
PR 11-MAR-1999; 99US-0122566P.
PR 28-JAN-2000; 2000US-00494096.
PR 16-MAR-2001; 2001US-0276822P.
XX
PA (CAPL/) CAPLAN M J.
PA (SOSI/) SOSIN H B.
PA (SAMP/) SAMPSON H.
PA (BARN/) BARNON G A.
PA (BURK/) BURKS A W.
PA (COCK/) COCKRELL G.
PA (COMP/) COMPADRE C M.
PA (CONN/) CONNAUGHTON C.
PA (HELM/) HELM R M.
PA (KING/) KING N E.
PA (KOPP/) KOPPER R A.
PA (MALE/) MALEKI S J.

PA (RABJ/) RABJOHN P A.
PA (SHIN/) SHIN D S.
PA (STAN/) STANLEY J S.
XX
XX
PI Caplan MJ, Sosin HB, Sampson H, Bannon GA, Burks AW, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabjohn PA, Shin DS, Stanley JS;
XX
DR WPI: 2003-875632/81.
DR N-PSDB; ADG27517.
XX
PT New modified anaphylactic food allergen comprising a cysteine residue
PT which has been modified so that it cannot participate in the disulfide
PT bond, useful for treating allergic reactions or wounds.
XX
PS Claim 27; SEQ ID NO 63; 194pp; English.
XX
CC The invention relates to a modified anaphylactic food allergen whose
CC amino acid sequence is substantially identical to that of a natural
CC anaphylactic food allergen. The natural anaphylactic food allergen
CC includes at least one cysteine residue that participates in a disulphide
CC bond when the natural anaphylactic food allergen is in its native
CC conformation, except that the cysteine residue has been modified so that
CC it cannot participate in the disulphide bond. Also included are a method
CC of making a modified anaphylactic food allergen, a nucleotide molecule
CC encoding a modified anaphylactic food allergen defined above, a
CC nucleotide molecule for causing a site specific mutation in a gene
CC encoding a natural anaphylactic food allergen, a transgenic plant or
CC animal expressing a modified anaphylactic food allergen defined above, a
CC method of treating an individual by reducing the clinical response to a
CC natural anaphylactic food allergen by administering a modified
CC anaphylactic food allergen and an isolated fragment of peanut allergen
CC Ara h 1, comprising at least 10 consecutive amino acids of ADG27464 or
CC ADG27465. About 10-17% of the amino acids have been modified in at least
CC one IgE epitope or all the IgE epitopes recognised when the natural
CC anaphylactic food allergen is contacted with serum IgE from individual(s)
CC allergic to the natural anaphylactic food allergen. The invention
CC discloses Peanut allergens Ara h1, Ara h2, Ara h3 (and their encoding
CC cDNAs), Soybean glycinin A2B1a and IgE-binding epitopes of the English
CC walnut allergen Jug n1. The modified anaphylactic food allergen can be
CC used for treating allergic reactions or wounds. The present sequence
CC represents a Peanut allergen of the invention (or its fragment).
XX
SQ Sequence 157 AA;
Alignment Scores:
Pred. No.: 9,328-90 Length: 157
Score: 842.00 Matches: 157
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 65.7% Indels: 0
DB: 7 Gaps: 0
US-10-728-323-2 (1-717) x ADG27518 (1-157)
QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTCGTCCCGCCATCTCGAGGCAG 61
Db 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisSerIleArgGln 20
QY 62 CAGTGGGAACCTCCAAGGAGACAGAGATGCCAGAGCCAGCTCGAGGGCGACCTGAGG 121
Db 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCCTCGGAGCAACATCTCATCGAGAGATCCACCTGACGAGGATTCATATGAACGGNC 181
Db 41 ProCysGluGlnHisLeuMetGlnIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGA 241
Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
QY 242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTCAGAACACCAAGG 301

Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100
QY 302 TGCATGTGCAGGCAATTGCACAGATCATGGAAACAGAGCCGATAGGTTGCAGGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CACACGAGCAACAGTTCCAGAGGGAGCTCAGGAACCTTCGCTCAACAGTGCAGCCTTAGG 421
Db 121 GlnGlnGlnGlnGlnPheLysArgGluLeuArgAsnGlnProGlnGlnCysGlyLeuArg 140
QY 422 GCACCCAGCGTTGCCACTTGGACGTCGAAAGTGGGGCAGAGACAGATAC 472
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157
RESULT 7
ADV97542
ID ADV97542 standard; protein; 157 AA.
XX
AC ADV97542;
XX
DT 10-MAR-2005 (first entry)
XX
DE Peanut Ara h II allergen (clone P38) seed storage protein.
XX
KW allergen; atopic dermatitis; diagnostic; immunotherapy;
KW immediate type hypersensitivity; Ara h II; dermatological; antiallergic;
KW immunosuppressive; antibody engineering; seed storage protein.
XX
OS Arachis hypogaea.
XX
PN US6835824-B1.
XX
PD 28-DEC-2004.
XX
PF 13-NOV-1998; 98US-00191593.
XX
PR 29-DEC-1995; 95US-0009455P.
PR 23-SEP-1996; 96US-00717933.
XX
PA (UYAR-) UNIV ARKANSAS.
XX
PI Burks AW, Stanley JS, Bannon GA, Cockrell G, Helm RW;
XX
DR WPI; 2005-045982/05.
DR N-PSDB; ADV97538; ADV97541.
XX
PT New isolated nucleic acid encoding peanut allergen Ara h 1, useful for
PT obtaining diagnostics with optimal concentrations of each allergen, for
PT developing panels of mixtures of recombinant allergens, and in
PT immunotherapy.
XX
PS Disclosure; SEQ ID NO 10; 141pp; English.
XX
CC This invention relates to a novel isolated nucleic acid molecule encoding
CC the peanut allergen Ara h I that has been identified in patients with
CC atopic dermatitis. Specifically, it refers to a monoclonal antibody
CC enzyme-linked immunosorbent assay (ELISA) used to detect and quantify the
CC peanut allergen, where the monoclonal antibodies have specificity for a
CC selected peanut allergen antigen. The present invention also describes
CC hybridomas that produce two-site monoclonal antibodies specific for a
CC peanut allergen that can be used in an ELISA to detect and determine the
CC concentration of a specific peanut allergen in a food product or in food
CC processing. Furthermore, it provides an expression vector and the
CC isolated nucleic acid or its fraction can include a diagnostic label.
CC Accordingly, these nucleic acids are useful for obtaining diagnostics
CC with optimal concentrations of each allergen or for developing panels of
CC mixtures of large numbers of recombinant allergens such that it can be
CC used in immunotherapy for the treatment of food hypersensitivity
CC reactions. This polypeptide sequence is the peanut Ara h II allergen
CC protein (p38 clone) of the invention.
XX
SQ Sequence 157 AA;

Alignment Scores: 9.32e-90 Length: 157
Pred. No.: 842.00 Matches: 157
Score: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 65.7% Indels: 0
DB: 9 Gaps: 0
US-10-728-323-2 (1-717) x ADV97542 (1-157)
QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTCGTCCACGATCTCGAGGCAG 61
Db 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAATCTCAAGGACACAGAGATCCAGAGCCAGCTCGAGGGCGACCTGAGG 121
Db 21 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCCTGCGAGCAACATCTCATGTCAGAAAGATCCAAACGTGACGAGGATTCATATGAACGGGAC 181
Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCTGGA 241
Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 80
QY 242 TCCTCTCAGACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAACACCAAGG 301
Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100
QY 302 TGCATGTGCAGGCATTGCAACAGATCATGGAAACAGAGCATAGGTTGCAGGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGGACCAACAGTTCAAGAGGGAGCTCAGAACTTCGCTCAACAGTGCAGCCTTAGG 421
Db 121 GlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACCCAGCGTTGCCACTTGGACGTCGAAAGTGGGGCAGAGACAGATAC 472
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157
RESULT 8
ADV97603
ID ADV97603 standard; protein; 159 AA.
XX
AC ADV97603;
XX
DT 10-MAR-2005 (first entry)
XX
DE Peanut Ara h II allergen (clone P38) seed storage protein.
XX
KW allergen; atopic dermatitis; diagnostic; immunotherapy;
KW immediate type hypersensitivity; Ara h II; dermatological; antiallergic;
KW immunosuppressive; antibody engineering; seed storage protein.
XX
OS Arachis hypogaea.
XX
PN US6835824-B1.
XX
PD 28-DEC-2004.
XX
PF 13-NOV-1998; 98US-00191593.
XX
PR 29-DEC-1995; 95US-0009455P.
PR 23-SEP-1996; 96US-00717933.
XX
PA (UYAR-) UNIV ARKANSAS.
XX
PI Burks AW, Stanley JS, Bannon GA, Cockrell G, Helm RW;
XX
DR WPI; 2005-045982/05.
DR N-PSDB; ADV97602.

XX New isolated nucleic acid encoding peanut allergen Ara h 1, useful for
PT obtaining diagnostics with optimal concentrations of each allergen, for
PT developing panels of mixtures of recombinant allergens, and in
PT immunotherapy.
XX
XX Disclosure; Fig 27; 141pp; English.
XX
XX This invention relates to a novel isolated nucleic acid molecule encoding
CC the peanut allergen Ara h I that has been identified in patients with
CC atopic dermatitis. Specifically, it refers to a monoclonal antibody
CC enzyme-linked immunosorbent assay (ELISA) used to detect and quantify the
CC peanut allergens, where the monoclonal antibodies have specificity for a
CC selected peanut allergen antigen. The present invention also describes
CC hybridomas that produce two-site monoclonal antibodies specific for a
CC peanut allergen that can be used in an ELISA to detect and determine the
CC concentration of a specific peanut allergen in a food product or in food
CC processing. Furthermore, it provides an expression vector and the
CC isolated nucleic acid or its fraction can include a diagnostic label.
CC Accordingly, these nucleic acids are useful for obtaining diagnostics
CC with optimal concentrations of each allergen or for developing panels of
CC mixtures of large numbers of recombinant allergens such that it can be
CC used in immunotherapy for the treatment of food hypersensitivity
CC reactions. This polypeptide sequence is the peanut Ara h II allergen
CC protein (P38 clone) of the invention.
XX
XX Sequence 159 AA;

Alignment Scores:
Pred. No.: 9.37e-90 Length: 159
Score: 842.00 Matches: 157
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 65.7% Indels: 0
DB: 9 Gaps: 0

US-10-728-323-2 (1-717) x ADV97603 (1-159)

QY 2 CTCACCATAGTACTAGTCCCTCGCCCTTTCTCTCGCTGCCACGCATCTCGAGGCAG 61
DB 3 LeuThriLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 22
QY 62 CAGTGGGAATCTCAAGGAGACAGAAGATCCAGAGCCAGCTCGAGAGCGCGAACCTTGAGG 121
DB 23 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 42
QY 122 CCTCGCAGCAATCTCATGACAGATCCACGTCGACGAGGATTCATATGAACGGAC 181
DB 43 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 62
QY 182 CGTACAGCCCTAGTCAGGATCGGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 241
DB 63 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 82
QY 242 TCCTCTCAGCACCAAGAGAGGTGTTCGAATGAGCTGAACGAGTTTGAGAACCAACCAAGG 301
DB 83 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 102
QY 302 TGCATGTGGAGCATTTGCACAGATCATGAGAACCAAGAGCATGTTGAGGGGAGG 361
DB 103 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 122
QY 362 CAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTCGCTCAACAGTGGCGCTTAGG 421
DB 123 GlnGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 142
QY 422 GCACCAAGCGTTTGCAGCTTGGACGTGGAAGTGGCGGCGAGACAGACATATAC 472
DB 143 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 159

RESULT 9
AAB82383
ID AAB82383 standard; protein; 207 AA.

XX AAB82383;
AC 11-SEP-2003 (revised)
DT 23-JUL-2001 (first entry)
XX
XX Peanut allergen Ara h2 gene product.
XX
XX Peanut; allergen; Ara h2; transgenic plant; allergy.
XX Arachis hypogaea.
XX Key Location/Qualifiers
FH Peptide 1..21
FT /label= Signal_peptide
FT Protein 22..207
FT /label= Mature_protein
XX WO200136621-A2.
XX
XX 25-MAY-2001.
XX 20-NOV-2000; 2000WO-US031657.
XX 19-NOV-1999; 99US-0167255P.
XX (UYAL-) UNIV ALABAMA A & M.
XX
XX Dodo HW, Arntzen CJ, Konan KN, Viquez OM;
XX WPI; 2001-355630/37.
DR N-PSDB; AAF90336, AAF90337.
XX
XX Producing transgenic peanut plants that produce allergen-free seeds,
PT useful in non-allergenic foods, by antisense or sense co-suppression of
PT allergen-encoding genes.
XX
XX Example 1; Fig 2; 72pp; English.
XX
XX The present sequence is that of the peanut allergenic protein (AP)
CC encoded by the Ara h2 gene (see AAF90336). The invention relates to a
CC method for producing a peanut plant having reduced, or undetectable, AP
CC content in its seed. A peanut plant cell is transformed with a DNA
CC construct containing an antisense AP gene and/or sense AP gene, or their
CC fragments, regenerated to plants, and fertile transgenic plants that
CC produce seeds with reduced AP content are identified. The AP sense or
CC antisense gene may comprise at least a portion of the Ara h2 gene
CC sequence. The seeds are useful for preparation of allergen-free foods.
CC Recombinant AP may be produced and used to produce antibodies useful for
CC detecting AP in foods, and for treatment or prevention of peanut allergy.
CC (Updated on 11-SEP-2003 to standardise OS field)
XX
XX Sequence 207 AA;

Alignment Scores:
Pred. No.: 2.71e-89 Length: 207
Score: 838.50 Matches: 162
Percent Similarity: 90.3% Conservative: 5
Best Local Similarity: 87.6% Mismatches: 17
Query Match: 65.5% Indels: 1
DB: 4 Gaps: 1

US-10-728-323-2 (1-717) x AAB82383 (1-207)

QY 2 CTCACCATAGTACTAGTCCCTCGCCCTTTCTCTCGCTGCCACGCATCTCGAGGCAG 61
DB 4 LeuThriLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 23
QY 62 CAGTGGGAATCTCAAGGAGACAGAAGATCCAGAGCCAGCTCGAGAGCGCGAACCTTGAGG 121
DB 24 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
QY 122 CCTCGCAGCAATCTCATGACAGATCCACGTCGACGAGGATTCATATGAACGGAC 181

|||||
Db 44 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 63
QY 182 CCCTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGGCGCTGGA 241
Db 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 83
QY 242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAACACCAAGG 301
Db 84 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 103
QY 302 TGCATGTGCGAGCGATTGCAACAGATCATGAGAACAGAGCCGATAGGTTGCGAGGGAGG 361
Db 104 CysMetCysGlnAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123
QY 362 CAACAGAGCAACAGATTCAAGAGGAGCTCAGGAACCTTGCTCAACAGTGCAGCCTTAGG 421
Db 124 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143
QY 422 GCACCAAGCGTTGCGACTTGGAGCTGCGAAAGTGGCGGCGAGACAGATACTAAACACCT 481
Db 144 AlaProGlnArgCysAspLeuAspValGluSerGlyArgArgProArgIleProPro 163
QY 482 ATCTCAAAAGAAAGAAAGAAAGAAAGAAATAGCTTATATATAAGC---TATTATCTA 538
Db 164 IleLeuThrGlySerArgSerArgArgHisGlnSerProTyrGlyAsnArgArgTyrSer 183
QY 539 TGGTTATGTTTAGTT 553
Db 184 AlaMetCysLeuLeu 188
RESULT 10
AAV40973
ID AAV40973 standard; protein; 156 AA.
XX AC AAV40973;
XX 17-OCT-2003 (revised)
DT 06-DEC-1999 (first entry)
XX 11-MAR-1999; 99US-0007763.
DE Ara h 2 protein fragment.
XX Peanut; allergen; Ara h 1; IgE; immunoglobulin E; epitope; Ara h 3;
KW allergic reaction; Ara h 2.
XX Arachis hypogaea.
XX WO945961-A1.
XX 16-SEP-1999.
XX 12-MAR-1999; 99WO-US005494.
XX 12-MAR-1998; 98US-0077763P.
PR 11-MAR-1999; 99US-0007763.
XX (UYAR-) UNIV ARKANSAS.
XX Burks W, Helm RM, Cockrell G, Bannon GA, Stanley JS, Shin DS;
PI Sampson H, Compadre CM, Huang SK, Maleki SJ, Kopper RA;
XX WPI; 1999-551218/46.
XX Tertiary structure of peanut allergen Ara h 1 for protection of a host
PT animal from allergic reaction.
XX Disclosure; Page 104; 193pp; English.
XX The invention provides a tertiary structure for the peanut allergen Ara H
CC 1. The Ara H 1 allergen is found to contain 23 linear IgE-binding
CC epitopes. The invention also provides an isolated recombinant peanut
CC allergen designated Ara h 3 and a nucleotide molecule encoding the peanut
CC allergen Ara h 3. Molecules of the invention are used to protect a host

CC animal from allergic reaction, particularly using a modified allergen
CC which is less reactive with IgE. The invention may also be used to ensure
CC that the allergen is not introduced into genetically modified food. The
CC present sequence represents an Ara h 2 protein fragment. (Updated on 17-
CC OCT-2003 to standardise OS field)
XX Sequence 156 AA;
SQ
Alignment Scores: 2.42e-88 Length: 156
Pred. No.: 830.00 Matches: 155
Score: 830.00
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 64.8% Indels: 0
DB: Gaps: 0
US-10-728-323-2 (1-717) x AAV40973 (1-156)
QY 2 CTCACCATAGTAGCCCTCGCCCTTTTCCTCTCGCTGCCACGCATCTGCGAGGCGAG 61
Db 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAATCTCAAGAGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAAACCTGAGG 121
Db 21 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCTCGGAGCAACATCTCATGCGAAGATCCAACTGAGAGGATTCATATGACGGGAC 181
Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGGCGCTGGA 241
Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
QY 242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAACACCAAGG 301
Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100
QY 302 TGCATGTGCGAGCGATTGCAACAGATCATGAGAACAGAGCCGATAGGTTGCGAGGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGAGCAACAGATTCAAGAGGAGCTCAGGAACCTTGCTCAACAGTGCAGCCTTAGG 421
Db 121 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACCAAGCGTTGCGACTTGGAGCTGCGAAAGTGGCGGCGAGAGAC 466
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAsp 155
RESULT 11
ADV97553
ID ADV97553 standard; protein; 157 AA.
XX AC ADV97553;
XX 10-MAR-2005 (first entry)
XX Peanut Ara h II allergen (clone P38) deduced protein sequence.
DE
XX
XX allergen; atopic dermatitis; diagnostic; immunotherapy;
KW immediate type hypersensitivity; Ara h II; dermatological; antiallergic;
KW immunosuppressive; antibody engineering; seed storage protein.
XX Arachis hypogaea.
XX OS
XX US6835824-B1.
XX 28-DEC-2004.
XX 13-NOV-1998; 98US-00191593.
XX 29-DEC-1995; 95US-0009455P.
PR

PR 23-SEP-1996; 96US-00717933.
XX (UYAR-) UNIV ARKANSAS.
XX Burks AW, Stanley JS, Bannon GA, Cockrell G, Helm RM;
XX WPI; 2005-045982/05.
XX
XX New isolated nucleic acid encoding peanut allergen Ara h 1, useful for
XX PT obtaining diagnostics with optimal concentrations of each allergen, for
XX PT developing panels of mixtures of recombinant allergens, and in
XX PT immunotherapy.
XX
XX Disclosure; SEQ ID NO 21; 141pp; English.
XX
XX This invention relates to a novel isolated nucleic acid molecule encoding
XX CC the peanut allergen Ara h I that has been identified in patients with
XX CC atopic dermatitis. Specifically, it refers to a monoclonal antibody
XX CC enzyme-linked immunosorbent assay (ELISA) used to detect and quantify the
XX CC peanut allergens, where the monoclonal antibodies have specificity for a
XX CC selected peanut allergen antigen. The present invention also describes
XX CC hybridomas that produce two-site monoclonal antibodies specific for a
XX CC peanut allergen that can be used in an ELISA to detect and determine the
XX CC concentration of a specific peanut allergen in a food product or in food
XX CC processing. Furthermore, it provides an expression vector and the
XX CC isolated nucleic acid or its fraction can include a diagnostic label.
XX CC Accordingly, these nucleic acids are useful for obtaining diagnostics
XX CC with optimal concentrations of each allergen or for developing panels of
XX CC mixtures of large numbers of recombinant allergens such that it can be
XX CC used in immunotherapy for the treatment of food hypersensitivity
XX CC reactions. This polypeptide sequence is the deduced peanut Ara h II
XX CC allergen protein (P38 clone) of the invention.
XX
XX SQ Sequence 157 AA;

Alignment Scores:
Pred. No.: 5,48e-88 Length: 157
Score: 827.00 Matches: 155
Percent Similarity: 98.7% Conservative: 0
Best Local Similarity: 98.7% Mismatches: 2
Query Match: 64.6% Indels: 0
DB: 9 Gaps: 0

US-10-728-323-2 (1-717) x ADV97553 (1-157)

QY 2 CTCACCATAGTACCCCTCGCCCTTTCTCTCGTCCGACGATCGAGGCGAGG 61
Db 1 LeuThrIleuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAACCTCAAGGAGACAGAAGATCCAGAGCCAGCTCGAGAGGCGGAACCTGAGG 121
Db 21 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCTGGGAGCAATCTCATGAGAGATCCAACTGAGAGGATTCATATGAACGGGAC 181
Db 41 ProCysGluGlnHisLeuMetGlnIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CGGTACAGCCCTAGTCAGGATCGGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 241
Db 61 ProSerTyrProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
QY 242 TCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGGAACGAGTTTGAGAACCAACCAAGG 301
Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100
QY 302 TGCATGTGCGAGCATTCGACAGATCATGGAACACGAGGCGATGAGTTGCGAGGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGGACCAAGTTCAAGAGGGAGCTCAGGAACCTTCCTCAACAGTGGCGGCTTAGG 421
Db 121 GlnGlnGluGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140

QY 422 GCACACACAGCGTTGCGACTTGACCTCGAAAGTGGCGGCGAGACAGATAC 472
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157
RESULT 12
ADO38314
ID ADO38314 standard; peptide; 156 AA.
XX
XX ADO38314;
AC
XX 15-JUL-2004 (first entry)
DT
XX
XX Ara h 2 MHC-class II-presented epitope #1.
DE
XX
XX Antibacterial; Virucide; Fungicide; Antiparasitic; Antiarthritic;
KW Antirheumatic; Neuroprotective; Antiinflammatory; Dermatological;
KW Immunosuppressive; Antidiabetic; Antithyroid; Antiasthmatic;
KW Antiallergic; Cystostatic; Antipsoriatic; Gene therapy; Vaccine;
KW MHC Class II; Ii-key motif; immune response; anthrax; EBOLA; HIV;
KW influenza; vaccinia virus; infection; bacterium; virus; parasite; fungus;
KW rickettsia; rheumatoid arthritis; multiple sclerosis;
KW lupus erythematosus; diabetes mellitus; myasthenia gravis;
KW autoimmune thyroiditis; scleroderma; dermatomyositis; pemphigus; asthma;
KW allergic rhinitis; topical dermatitis; colitis; cancer; psoriasis;
KW adenoma; peanut; Ara h 2.
XX
XX Arachis hypogaea.
OS
XX US2004058881-A1.
PN
XX 25-MAR-2004.
PD
XX
XX 24-SEP-2002; 2002US-00253286.
PF
XX
XX 24-SEP-2002; 2002US-00253286.
PR
XX
XX (ANTI-) ANTIGEN EXPRESS INC.
PA
XX Humphreys RE, Xu M;
PI
XX WPI; 2004-294259/27.
DR
XX
XX New non-naturally occurring protein or polypeptide modified by
PT recombinant DNA techniques, useful for treating multiple sclerosis,
PT diabetes mellitus, myasthenia gravis, scleroderma, allergic rhinitis,
PT colitis, cancer or psoriasis.
XX
XX Example 2; Page 18; 90pp; English.

XX The invention relates to a non-naturally occurring protein or polypeptide
CC (i) modified by recombinant DNA techniques comprising: a C-terminal
CC element comprising an MHC Class II-presented epitope; an N-terminal
CC element comprising an Ii-key motif; and an intervening element comprising
CC a sequence of 4-11 amino acid residues where the modification by
CC recombinant DNA techniques taking place within elements (b) and (c). Also
CC described are methods for: suppressing or enhancing an immune response
CC directed toward an MHC (major histocompatibility complex) Class II-
CC presented epitope of interest. Suppressing an immune response directed
CC toward an MHC Class II-presented epitope of interest comprises: providing
CC a nucleic acid sequence encoding the MHC Class II-presented epitope of
CC interest, the nucleic acid sequence encoding an Ii-key motif located 4-11
CC amino acids upstream from the N-terminal residue of the MHC Class II-
CC presented epitope of interest; and modifying the Ii-key motif to decrease
CC its conformance to the archetypal Ii-key regulatory motif. Enhancing an
CC immune response directed toward an MHC Class II-presented epitope of
CC interest comprises: providing a nucleic acid sequence encoding the MHC
CC Class II-presented epitope of interest, the nucleic acid sequence lacking
CC an Ii-key motif located 4-11 amino acids upstream from the N-terminal
CC residue of the MHC Class II-presented epitope of interest; and modifying
CC the nucleic acid sequence to introduce an Ii-key motif appropriately
CC spaced from the MHC Class II-presented epitope. The protein or
CC polypeptide of interest corresponds to a protein or polypeptide encoded
CC by an infectious pathogen selected from anthrax, EBOLA, HIV or influenza,

CC preferably vaccinia virus. The non-naturally occurring protein or
 CC polypeptide (1) modified by recombinant DNA techniques is useful for
 CC treating infectious diseases caused or associated with infection by a
 CC bacterium, virus, parasite, fungus, rickettsia or other infectious
 CC agents. It is also useful for treating rheumatoid arthritis, multiple
 CC sclerosis, lupus erythematosus, diabetes mellitus, myasthenia gravis,
 CC autoimmune thyroiditis, scleroderma, dermatomyositis, pemphigus, asthma,
 CC allergic rhinitis, topical dermatitis, colitis, cancer, psoriasis or
 CC adenomas. The present sequence represents the amino acid sequence of a
 CC peanut allergen Ara h 2 MHC class II-presented epitope used in the
 CC invention.

XX SQ Sequence 156 AA;

Alignment Scores:
 Pred. No.: 4.8e-87 Length: 156
 Score: 819.00 Matches: 153
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 63.9% Indels: 0
 DB: 8 Gaps: 0

US-10-728-323-2 (1-717) x ADO38314 (1-156)

QY 2 CTCACATAGTACGCTCGCCCTTTCTCTCGTCCACGATCTGCGAGCAG 61
 DB 4 LeuThrIleuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 23
 QY 62 CAGTGGGAATCCAAAGGAGACAGAAGATGCCAGCGAGCTCCAGAGGGGGAACCTCAGG 121
 DB 24 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
 QY 122 CCTTGGAGCAACATCTCATCGAAGATCCAACTGACGAGGATTCATATGAACGGGAC 181
 DB 44 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 63
 QY 182 CCGTACAGCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGCGCTTGA 241
 DB 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 83
 QY 242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAACCAACCAAGG 301
 DB 84 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArg 103
 QY 302 TGCATGTCCAGGATTCGCAACAGATCATGAGAACGAGCGCATAGGTTGCGGGGAGG 361
 DB 104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123
 QY 362 CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTCGGCCTTAGG 421
 DB 124 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143
 QY 422 GCACACAGCGTTGCGACTTGAGCGTTCGAAAGTGGCGGC 460
 DB 144 AlaProGlnArgCysAspLeuAspValGluSerGlyGly 156

RESULT 13

ID ABU52576 standard; protein; 157 AA.

XX AC ABU52576;

XX 10-MAR-2003 (first entry)

DE Peanut Ara h2 mutant 5 (Q20A/Q31A/D60A/D67A).

XX Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;
 KW mutein; anaphylactic food allergen; antiallergenic; vaccine;
 KW wound healing.

XX OS Homo sapiens.

XX WO200274250-A2.

PN

XX 26-SEP-2002.

XX 18-MAR-2002; 2002WO-US009108.

XX 16-MAR-2001; 2001US-0276822P.

XX 18-MAR-2002; 2002US-00276822.

XX (PANA-) PANACEA PHARM.

XX Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
 PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
 PI Rabin PA, Shin DS, Stanley JS;

XX WPI; 2003-018765/01.

XX New modified anaphylactic food allergen, useful for preventing or
 PT treating allergic reactions associated with e.g. anaphylactic allergens.

XX Example 13; Page: 300pp; English.

XX The invention relates to a modified anaphylactic food allergen has an
 CC amino acid sequence that is substantially identical to that of natural
 CC anaphylactic food allergen, except for a cysteine residue that has been
 CC modified so that it cannot participate in the disulphide bond. The
 CC modification may also comprise mutation of the IgE binding sites to
 CC reduce allergenicity. Also included are: (1) a method of making a
 CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
 CC or for causing a site specific mutation in the modified anaphylactic food
 CC allergen; (3) a transgenic plant or animal expressing the modified
 CC anaphylactic food allergen; (4) a method of treating an individual by
 CC reducing the clinical response to a natural anaphylactic food allergen;
 CC and an isolated fragment of peanut allergen Ara h 1. The modified
 CC anaphylactic food allergen is useful for preventing or treating allergic
 CC reactions associated with any natural allergen such as food, insect,
 CC rubber or preferably anaphylactic allergens. It is also useful for
 CC treating wounds in mammals such as bovine, canine, feline, caprine,
 CC ovine, porcine, murine or equine species. The present sequence is a food
 CC allergen, mutated to alter its IgE binding characteristics. Note: The
 CC present sequence is not shown in the specification but was created by the
 CC indexer using information provided in the specification

XX Sequence 157 AA;

Alignment Scores:

Pred. No.: 1.87e-86 Length: 157
 Score: 814.00 Matches: 153
 Percent Similarity: 97.5% Conservative: 0
 Best Local Similarity: 97.5% Mismatches: 4
 Query Match: 63.5% Indels: 0
 DB: 6 Gaps: 0

US-10-728-323-2 (1-717) x ABU52576 (1-157)

QY 2 CTCACATAGTACGCTCGCCCTTTCTCTCGTCCACGATCTGCGAGCAG 61

DB 1 LeuThrIleuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgAla 20

QY 62 CAGTGGGAATCCAAAGGAGACAGAAGATGCCAGCGAGCTCCAGAGGGGGAACCTCAGG 121

DB 21 GlnTrpGluLeuGlnGlyAspArgArgCysAlaSerGlnLeuGluArgAlaAsnLeuArg 40

QY 122 CCTTGGAGCAACATCTCATCGAAGATCCAACTGACGAGGATTCATATGAACGGGAC 181

DB 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAla 60

QY 182 CCGTACAGCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGCGGCTTGA 241

DB 61 ProTyrSerProSerGlnAlaProTyrSerProSerProTyrAspArgGlyAlaGly 80

QY 242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAACCAACCAAGG 301

DB 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArg 100

QY 302 TGCATGTGAGGCGATTGCAACAGATCATGGAGACCGAGCGGATAGTTGCGAGGCGG 361
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTCCCTCAACAGTGCAGCGCTTAGG 421
Db 121 GlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACACACAGGTTGCCACTTGGACGTCGAAAGTGGCGGCGAGACAGATAC 472
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157
RESULT 14
ADM12096
ID ADM12096 standard; protein; 156 AA.
XX
AC ADM12096;
XX
DT 20-MAY-2004 (first entry)
XX
DE Arachis hypogaea 2 (Ara h2) protein.
XX
KW antigen presentation enhancing hybrid polypeptide; mammalian II-Key;
KW MHC class II; antibacterial; virucide; fungicide; antirheumatic;
KW antiarthritic; neuroprotective; dermatological; immunosuppressive;
KW antiinflammatory; antidiabetic; antithyroid; immune;
KW rheumatoid arthritis; multiple sclerosis; lupus erythematosus;
KW diabetes mellitus; myasthenia gravis; autoimmune thyroiditis;
KW scleroderma; dermatomyositis; pemphigus.
XX
OS Arachis hypogaea.
XX
PN US2003235594-A1.
XX
PD 25-DEC-2003.
XX
PF 17-SEP-2002; 2002US-00245871.
XX
PR 14-SEP-1999; 99US-00396813.
XX
PR 17-JUL-2002; 2002US-00197000.
XX
PA (ANTI-) ANTIGEN EXPRESS INC.
XX
XX Humphreys R, Xu M;
XX WPI; 2004-070554/07.
XX
PT Novel II-Key/antigen presentation enhancing hybrid polypeptide, useful
PT for treating infections, rheumatoid arthritis, multiple sclerosis, lupus
PT erythematosus and diabetes mellitus.
XX
PS Example 2; Page 19; 87pp; English.
XX
CC The invention relates to a novel antigen presentation enhancing hybrid
CC polypeptide. The novel polypeptide has an N-terminal element consisting
CC of 4-16 residues of a mammalian II-Key peptide and its non-N-terminal
CC deletion modifications, a chemical structure covalently linking the N-
CC terminal element to an MHC class II-presented epitope of a C-terminal
CC element. The C-terminal element comprises an antigenic epitope, which
CC binds to an antigenic peptide binding site of an MHC class II molecule.
CC The antigen presentation enhancing hybrid polypeptide has the following
CC activities: antibacterial, virucide, fungicide, antirheumatic,
CC antiarthritic, neuroprotective, dermatological, immunosuppressive,
CC antiinflammatory, antidiabetic, and antithyroid. The antigen presentation
CC enhancing hybrid polypeptide is useful for modulating the immune response
CC in an individual and for treating infections (such as bacteria, virus,
CC parasite and fungus), rheumatoid arthritis, multiple sclerosis, lupus
CC erythematosus, diabetes mellitus, myasthenia gravis, autoimmune
CC thyroiditis, scleroderma, dermatomyositis and pemphigus. This sequence
XX represents a mammalian II key related protein of the invention.
XX
SQ Sequence 156 AA;

Alignment Scores:
Pred. No.: 3 21e-86 Length: 156
Score: 812.00 Matches: 152
Percent Similarity: 99.3% Conservatives: 0
Best Local Similarity: 99.3% Mismatches: 1
Query Match: 63.4% Indels: 0
DB: 8 Gaps: 0
US-10-728-323-2 (1-1717) x ADM12096 (1-156)
QY 2 CTCACACATAGTAGCCCTCCCTTTCTCTCTCGTCCACGCGCATCTCGAGGCGAG 61
Db 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 23
QY 62 CAGTGGGAATCCCAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGCGCAACCTGAGG 121
Db 24 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
QY 122 CCCTGCGAGCAACATCTCATGCGAAGATCCCAACGTCGAGGAGTTTCATATGAACGGGAC 181
Db 44 ProCysGluGlyHisLeuMetGlnLysIleGlnAArgAspGluAspSerTyrGluArgAsp 63
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCGTGA 241
Db 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 83
QY 242 TCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGACGAGTTTGAGAACACCAAGG 301
Db 84 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnGlnArg 103
QY 302 TGCATGTGCGAGCGATTGCAACAGATCATGGAGAACCGAGCGATAGTTGCGAGGCGAGG 361
Db 104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123
QY 362 CAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCCTCAACAGTGCAGCGCTTAGG 421
Db 124 GlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143
QY 422 GCACACACAGCGTTGCGAGCTTGGACCTCGAAAGTGGCGGC 460
Db 144 AlaProGlnArgCysAspLeuAspValGluSerGlyGly 156
RESULT 15
ABU52575
ID ABU52575 standard; protein; 157 AA.
XX
AC ABU52575;
XX
DT 10-MAR-2003 (first entry)
XX
DE Peanut Ara h2 mutant 4 (E35A/P41A/D60A/D67A).
XX
KW Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;
KW mutein; anaphylactic food allergen; antiallergenic; vaccine;
XX wound healing.
XX
OS Homo sapiens.
XX
PN WO200274250-A2.
XX
PD 26-SEP-2002.
XX
PF 18-MAR-2002; 2002WO-US0009108.
XX
PR 16-MAR-2001; 2001US-0276822P.
XX
PR 18-MAR-2002; 2002US-00276822.
XX
PA (PANA-) PANACEA PHARM.
XX
PI Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabjohn PA, Shin DS, Stanley JG;

XX DR

WPI; 2003-018765/01.

XX

New modified anaphylactic food allergen, useful for preventing or treating allergic reactions associated with e.g. anaphylactic allergens.

XX

Example 13: Page; 300pp; English.

XX

The invention relates to a modified anaphylactic food allergen has an amino acid sequence that is substantially identical to that of natural anaphylactic food allergen, except for a cysteine residue that has been modified so that it cannot participate in the disulphide bond. The modification may also comprise mutation of the IGE binding sites to reduce allergenicity. Also included are: (1) a method of making a modified anaphylactic food allergen; (2) a nucleotide molecule encoding or for causing a site specific mutation in the modified anaphylactic food allergen; (3) a transgenic plant or animal expressing the modified anaphylactic food allergen; (4) a method of treating an individual by reducing the clinical response to a natural anaphylactic food allergen; and an isolated fragment of peanut allergen Ara h 1. The modified anaphylactic food allergen is useful for preventing or treating allergic reactions associated with any natural allergen such as food, insect, rubber or preferably anaphylactic allergens. It is also useful for treating wounds in mammals such as bovine, canine, feline, caprine, ovine, porcine, murine or equine species. The present sequence is a food allergen, mutated to alter its IGE binding characteristics. Note: The present sequence is not shown in the specification but was created by the indexer using information provided in the specification

XX

Sequence 157 AA;

Alignment Scores:

Pred. No.:	3,228-86	Length:	157
Score:	812.00	Matches:	153
Percent Similarity:	97.5%	Conservative:	0
Best Local Similarity:	97.5%	Mismatches:	4
Query Match:	63.4%	Indels:	0
DB:	6	Gaps:	0

US-10-728-323-2 (1-717) x ABUS2575 (1-157)

Qy	2	CTCACCATACTAGTAGCCCTCGCCCTTTTCCTCCTCGCTGCCCAAGCATCTCGGAGGAG	61
Db	1	LeuthrilleLeuValaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln	20
Qy	62	CAGTGGGAACCTCAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGCGCAACCTGAGG	121
Db	21	GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuAlaArgAlaasnLeuArg	40
Qy	122	CCCTGCGAGCAACATCTCATGCAGAGAAGATCCAAACGTGACGAGGATTCATATGAACGGGAC	181
Db	41	AlaCysGluGlnHisLeuMetGlnIlysileGlnArgAspGluAspSerTyrGluArgAla	60
Qy	182	CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA	241
Db	61	ProTyrSerProSerGlnAlaProTyrSerProSerProTyrAspArgArgGlyAlaGly	80
Qy	242	TCCTCTCAGCAACCAAGAGAGGTGTTGCCATGAGCTGAACAGTTCGTAGAACCAACCAAGG	301
Db	81	SerSerGlnHisGlnGluArgCysCysasnGluLeuAsnGluPheGluAsnAsnGlnArg	100
Qy	302	TGCATGTGCGAGGCATTGCAACAGATCATGGAAGAACACAGAGGATAGGTTGCAGGGGAGG	361
Db	101	CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg	120
Qy	362	CAACAGGAGCAACAGTTCGAAGAGGAGCTCAGCAACTTCCTCCACAGTTCGGCCCTTAGG	421
Db	121	GlnGlnGluGlnGlnPheIysArgGluLeuArgasnLeuProGlnGlnCysGlyLeuArg	140
Qy	422	GCACCACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGCAGACAGACATAC	472
Db	141	AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr	157

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: May 15, 2006, 22:18:32 ; Search time 4.3124 Seconds
(without alignments)
4123.811 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 1281
Sequence: 1 gctcaccatactagtagcccc.....taaaagatcatgtttgtgt 717

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abs/ABSWEB_spool/US10728323/runat_15052006_172137_22463/app_query.fasta_1
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs02p -USER=US10728323 @CN1.1.131 @runat_15052006_172137_22463
-NCPU=6 -ICPU=3 -NO.MMAP -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5 COMB.pep.*
2: /cgn2_6/prodata/1/iaa/6 COMB.pep.*
3: /cgn2_6/prodata/1/iaa/H COMB.pep.*
4: /cgn2_6/prodata/1/iaa/PCTUS COMB.pep.*
5: /cgn2_6/prodata/1/iaa/RE COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	842	65.7	157	2	US-09-106-872A-22
2	842	65.7	157	2	US-09-191-593-10
3	842	65.7	159	2	US-09-106-872A-2
4	838.5	65.5	207	2	US-09-715-036-2
5	827	64.6	157	2	US-09-191-593-21
6	243	19.0	158	1	US-08-618-911-4
7	243	19.0	158	2	US-08-938-675A-2
8	243	19.0	158	2	US-09-531-727-2
9	241.5	18.9	155	1	US-08-618-911-2
10	229	17.9	158	1	US-08-618-911-6
11	156	12.2	28	2	US-09-191-593-53
12	156	12.2	158	1	US-08-670-186-4
13	157	157	157	2	US-09-106-872A-22
14	157	157	157	2	US-09-191-593-10
15	159	159	159	2	US-09-106-872A-2
16	838.5	65.5	207	2	US-09-715-036-2
17	827	64.6	157	2	US-09-191-593-21
18	243	19.0	158	1	US-08-618-911-4
19	243	19.0	158	2	US-08-938-675A-2
20	243	19.0	158	2	US-09-531-727-2
21	241.5	18.9	155	1	US-08-618-911-2
22	229	17.9	158	1	US-08-618-911-6
23	156	12.2	28	2	US-09-191-593-53
24	156	12.2	158	1	US-08-670-186-4

13	151	11.8	158	1	US-08-670-186-6	Sequence 6, Appli
14	139	10.9	155	1	US-08-670-186-2	Sequence 2, Appli
15	139	10.9	184	1	US-08-453-924-3	Sequence 3, Appli
16	138	10.8	28	2	US-09-191-593-58	Sequence 58, Appli
17	136.5	10.7	164	2	US-09-627-536-13	Sequence 13, Appli
18	129.5	10.1	150	2	US-10-442-174A-1	Sequence 1, Appli
19	118	9.2	30	2	US-09-106-872A-16	Sequence 16, Appli
20	118	9.2	30	2	US-09-191-593-57	Sequence 57, Appli
21	106.5	8.3	174	2	US-09-645-593-7	Sequence 7, Appli
22	106.5	8.3	1088	2	US-09-233-857-13	Sequence 13, Appli
23	105.5	8.2	1099	2	US-09-442-100-2	Sequence 2, Appli
24	105.5	8.2	1099	2	US-08-939-108-2	Sequence 2, Appli
25	105.5	8.2	1099	2	US-09-442-103-2	Sequence 2, Appli
26	102	8.0	316	2	US-08-098-327E-31	Sequence 31, Appli
27	102	8.0	316	2	US-08-462-625-31	Sequence 31, Appli
28	100	7.8	17	2	US-09-191-593-55	Sequence 55, Appli
29	100	7.8	17	2	US-09-191-593-61	Sequence 61, Appli
30	100	7.8	17	2	US-09-191-593-62	Sequence 62, Appli
31	100	7.8	224	2	US-09-270-767-44901	Sequence 44901, A
32	99.5	7.8	148	2	US-09-083-852A-4	Sequence 4, Appli
33	99.5	7.8	148	2	US-09-083-852A-6	Sequence 6, Appli
34	99.5	7.8	148	2	US-09-489-674B-6	Sequence 6, Appli
35	97	7.6	185	2	US-09-252-991A-23490	Sequence 23490, A
36	95	7.4	32	1	US-08-158-704-1	Sequence 1, Appli
37	95	7.4	32	1	US-08-610-424B-1	Sequence 1, Appli
38	95	7.4	32	2	US-09-336-463-1	Sequence 1, Appli
39	95	7.4	32	2	US-09-191-593-1	Sequence 1, Appli
40	95	7.4	152	2	US-09-252-991A-19284	Sequence 19284, A
41	94	7.3	43	2	US-09-303-814-1	Sequence 1, Appli
42	94	7.3	500	2	US-09-265-630-13	Sequence 13, Appli
43	93.5	7.3	415	2	US-09-252-991A-18252	Sequence 18252, A
44	93	7.3	537	2	US-09-252-991A-27024	Sequence 27024, A
45	92.5	7.2	788	1	US-08-918-914-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-106-872A-22
; Sequence 22, Application US/09106872A
; Patent No. 6486311
; GENERAL INFORMATION:
; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Stanley, J. Steven
; APPLICANT: Cockrell, Gael
; APPLICANT: King, Nina E.
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Helm, Ricki M.
; APPLICANT: Bannan, Gary A.
; TITLE OF INVENTION: Peanut Allergens and Methods
; FILE REFERENCE: HS 103 CIP
; CURRENT APPLICATION NUMBER: US/09/106,872A
; CURRENT FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: PCT/US96/15222
; PRIOR FILING DATE: 1996-09-23
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-09-106-872A-22

Alignment Scores:
Pred. No.: 2,528-86 Length: 157
Score: 842.00 Matches: 157
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 65.7% Indels: 0
DB: 2 Gaps: 0

US-10-728-323-2 (1-717) x US-09-106-872A-22 (1-157)

Db 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCTGGAGCAACATCTCATGAGAGATCCAAACGTGACGAGGATTTCATATCAACGGGAC 181
Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGGCGCTGGA 241
Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
QY 242 TCCTCTCAGCACACAGAGAGGTTCATATGAGCTGAACGAGTTTCAGAACACCAAGG 301
Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArg 100
QY 302 TGCATGTGCGAGGATTCGACAGATCATGAGAACCAAGAGCATAGGTTCAGGGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGGACCAACAGTTCAGAGGGAGCTCAGAACTTCGCTCAACAGTCCGCTTAGG 421
Db 121 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACACAGCGTTGCGACTTGACGCTCGAAAGTGGCGGCAGACAGATAC 472
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157
RESULT 3
US-09-106-872A-2
; Sequence 2, Application US/09106872A
; Patent No. 6486311
; GENERAL INFORMATION:
; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Stanley, J. Steven
; APPLICANT: Cockrell, Gael
; APPLICANT: King, Nina E.
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Helm, Ricki M.
; APPLICANT: Bannon, Gary A.
; TITLE OF INVENTION: Peanut Allergens and Methods
; FILE REFERENCE: HS 103 CIP
; CURRENT APPLICATION NUMBER: US/09/106,872A
; CURRENT FILING DATE: 1999-06-29
; PRIOR FILING DATE: PCT/US96/15222
; PRIOR FILING DATE: 1996-09-23
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Arachis hypogaea
; FEATURE:
; OTHER INFORMATION: Amino Acids 17-26 are Ara H 2 Binding Epitopes,
; OTHER INFORMATION: Peptide 1
; OTHER INFORMATION: Amino Acids 23-32 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 2
; OTHER INFORMATION: Amino Acids 29-38 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 3
; OTHER INFORMATION: Amino Acids 41-50 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 4
; OTHER INFORMATION: Amino Acids 51-60 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 5
; OTHER INFORMATION: Amino Acids 59-68 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 6
; OTHER INFORMATION: Amino Acids 67-76 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 7
; OTHER INFORMATION: Amino Acids 117-126 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 8
; OTHER INFORMATION: Amino Acids 129-138 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 9
; OTHER INFORMATION: Amino Acids 145-154 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 10
; OTHER INFORMATION: Amino Acids 13-22 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 1, Table 12

; OTHER INFORMATION: Amino Acids 27-36 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 2, Table 12
; OTHER INFORMATION: Amino Acids 59-68 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 3, Table 12
; OTHER INFORMATION: Amino Acids 81-90 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 4, Table 12
; OTHER INFORMATION: Amino Acids 91-100 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 5, Table 12
; OTHER INFORMATION: Amino Acids 145-154 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 6, Table 12
US-09-106-872A-2
Alignment Scores:
Pred. No.: 2,53e-86 Length: 159
Score: 842.00 Matches: 157
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 65.7% Indels: 0
DB: 2 Gaps: 0
US-10-728-323-2 (1-717) x US-09-106-872A-2 (1-159)
QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCCTCGCTGCCACGCATCTCGGGGAG 61
Db 3 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaHisAlaSerAlaArgGln 22
QY 62 CAGTGGGAATCCCAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGCGCAACCTGAGG 121
Db 23 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 42
QY 122 CCTCGAGCAACATCTCATGAGAGATCCAACTGACGAGGATTCATATGAACGGGAC 181
Db 43 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 62
QY 182 CGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGGCGCTGGA 241
Db 63 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 82
QY 242 TCCTCTCAGCACACAGAGAGGTGTTCATAGCTGAACGAGTTTGAGAACCAACCAAGG 301
Db 83 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArg 102
QY 302 TGCATGTGCGAGGATTCGACAGATCATGAGAACCAAGAGCATAGGTTCAGGGGAGG 361
Db 103 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 122
QY 362 CAACAGGACCAACAGTTCAGAGGGAGCTCAGAACTTCGCTCAACAGTCCGCTTAGG 421
Db 123 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 142
QY 422 GCACACAGCGTTGCGACTTGACGCTCGAAAGTGGCGGCAGACAGATAC 472
Db 143 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 159
RESULT 4
US-09-715-036-2
; Sequence 2, Application US/09715036
; Patent No. 6943010
; GENERAL INFORMATION:
; APPLICANT: DODO, HORTENSE W.
; APPLICANT: ARNTZEN, CHARLES J.
; APPLICANT: KONAN, KOFFI N'DA
; APPLICANT: VIOUEZ, OLGA
; TITLE OF INVENTION: DOWN-REGULATION AND SILENCING OF ALLERGEN GENES IN
; FILE REFERENCE: 072121/0104
; CURRENT APPLICATION NUMBER: US/09/715,036
; CURRENT FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/167,255
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2

```
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-09-715-036-2

Alignment Scores:      6.95e-86      Length:      207
Pred. No.:            838.50         Matches:    162
Score:                90.3%          Conservative: 5
Percent Similarity:   87.6%          Mismatches: 17
Best Local Similarity: 65.5%         Indels:     1
Query Match:         2               Gaps:      1
DB:

US-10-728-323-2 (1-717) x US-09-715-036-2 (1-207)

Qy  2 CTCACATAGTCTAGTCCCTCGCCCTTTCTCTCTCGCTCCACGATCTCGAGGCGAG 61
Db  4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaAArgGln 23
Qy  62 CAGTGGGAATCTCAAGAGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGGGAACCTGAGG 121
Db  24 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
Qy  122 CCCTGCGAGCAATCTCATGCAGAGATCCACGTCGACGAGGATTCATATGAACGGAC 181
Db  44 ProCysGlnGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluAArgasp 63
Qy  182 CCCTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGCGCTGGA 241
Db  64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 83
Qy  242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTGAGAACCAACAAAGG 301
Db  84 SerSerGlnHisGlnGluAArgCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 103
Qy  302 TGCATGTCCAGGATTCGACATGACATGAGAGACGAGCGATGTTGCGAGGGAGG 361
Db  104 CysMetCysGluAlaLeuGlnGlnIleMetGlnAsnGlnSerAspArgLeuGlnGlyArg 123
Qy  362 CAACAGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCCCTCAACAGTCGGCGCTTAGG 421
Db  124 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnCysGlyLeuArg 143
Qy  422 GCACACAGCGTTGCGAGCTTGGAGCTGCAAGTGGCGGCGAGACAGACATACAAACCT 481
Db  144 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgArgProArgIleProPro 163
Qy  482 ATCTCAAAAAAGAAAAAGAAAAAGAAAAAGAAATAGCTTATATATAAGC---TATTATCTA 538
Db  164 IleLeuThrGlySerArgSerArgArgHisGlnSerProTyrGlyAsnArgArgTyrSer 183
Qy  539 TGGTTATGTTTAGTT 553
Db  184 AlaMetCysLeuLeu 188

RESULT 5
US-09-191-593-21
; Sequence 21, Application US/09191593
; Patent No. 6835824
; GENERAL INFORMATION:
; APPLICANT: BURKS, A Wesley, HELM, Ricki M.
; APPLICANT: COCKRELL, Gael, STANLEY, J Steven,
; TITLE OF INVENTION: PEANUT ALLERGENS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Head, Johnson & Kachigian
; STREET: 112 W. Center St., Suite 230
; CITY: Fayetteville
; STATE: Arkansas AR
; COUNTRY: United States of America
; ZIP: 72701

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS 6.2
; SOFTWARE: Wordperfect 6.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/191,593
; FILING DATE: 13 NOVEMBER 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/717,933
; FILING DATE: 23 SEPTEMBER 1996
; APPLICATION NUMBER: US 07/998,377
; FILING DATE: 30 DECEMBER 1992
; APPLICATION NUMBER: US 08/158,704
; FILING DATE: 29 NOVEMBER 1993
; APPLICATION NUMBER: US 60/009,455
; FILING DATE: 29 DECEMBER 1995
; APPLICATION NUMBER: US 08/610,424
; FILING DATE: 04 MARCH 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ALEXANDER, DANIEL R
; REGISTRATION NUMBER: 32,604
; REFERENCE/DOCKET NUMBER: ARK00895601B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (501) 582-9111
; TELEFAX: (501) 521-4931
; TELEX: No. 6835824 applicable
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 157 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. 6835824 applicable
; TOPOLOGY: unknown
; MOLECULE TYPE: glycoprotein
; DESCRIPTION: identified as Ara h 2 P38 deduced
; DESCRIPTION: sequence from nucleotide sequence reading frame.
; HYPOTHETICAL: No
; ANTI-SENSE: No. 6835824 applicable
; FRAGMENT TYPE: No. 6835824 applicable
; ORIGINAL SOURCE:
; ORGANISM: Arachis hypogaea
; STRAIN: Florunner
; INDIVIDUAL ISOLATE: Ara h II
; DEVELOPMENTAL STAGE:
; HAPLOTYPE: No. 6835824 applicable
; TISSUE TYPE:
; CELL TYPE: No. 6835824 applicable
; CELL LINE: No. 6835824 applicable
; ORGANELLE: No. 6835824 applicable
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: No. 6835824 applicable
; MAP POSITION: No. 6835824 applicable
; UNITS: No. 6835824 applicable
; FEATURE:
; NAME/KEY: final check
; LOCATION: 4..480
; IDENTIFICATION METHOD: By agreement with
; IDENTIFICATION METHOD: protein information and established
; IDENTIFICATION METHOD: consensus sequence
; OTHER INFORMATION: Seed storage protein and
; OTHER INFORMATION: allergen
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
```



```
QY 5 ACCATCTAGTAGCCCTCGCCCTTTCTCTCCCTGCGTCCACGCATCTCGGAGGCAGCAG 64
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 5 ThrileLeuIleSerLeuLeuPheCysIle---AlaHisThrCysSerAlaSerLys 23
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 65 TGGGAACCTCAAGGACACAGACAGATGCCAGAGCCAGCTCGAGGGCGGAACTGAGGCC 124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 24 TrpGlnHisGlnAsp---SerCysArgLysGlnLeuLysGlyValAsnLeuThrPro 42
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 125 TGGGAGCAACATCTCATGACAGATCCAA-----CGTCAGGAGGATCATATGAACGG 178
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 43 CysGluLysHisIleMetGluLysIleGlnGlyArgGlyAspAspAspAspAsp 62
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 179 GACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGCGCGT 238
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 63 AspAspAsnHisIleLeuArgThrMetArgGlyArgIleAsnTyrlleArgLysLysGlu 82
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 239 GGATCCTCTCAGCACCAAG-----AGGTGTTGCAATGAGCTGAAC 280
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 83 GlyLysAspGluAspGluGluGluGluGlnMetGlnLysCysThrGluMetSer 102
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 281 GAGTTTGCAGAACCAAGGTGCTGCGAGGCATTGCAACAGATCATGGAGAACCGAG 340
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 103 GluLeuLys---SerProLysCysGlnCysLysAlaLeuGlnLysIleMetGluAsnGln 121
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 341 AGCGATAGTTGAGGGAGGCAACAGGACCAAGTTCAGAGGGAGCTCAGGAACCTTG 400
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 122 SerGluGluLeuGlu---GluLysGluAsnLysLysMetGluLysGluLeuMetAsnLeu 140
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 401 CCTCAACAGTGGCGCCTTAGGGCCACCACGCGTTGGCACTTG 442
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 141 AlaThrMetCysArgPheGlyProMetIleGlyCysAspLeu 154
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 11

```
US-09-191-593-53
; Sequence 53, Application US/09191593
; Patent No. 6835824
; GENERAL INFORMATION:
; APPLICANT: BURKS, A Wesley, HELM, Ricki M,
; APPLICANT: COCKRELL, Gael, STANLEY, J Steven,
; APPLICANT: BANNON, Gary A
; TITLE OF INVENTION: PEANUT ALLERGENS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Head, Johnson & Kachigian
; STREET: 112 W. Center St., Suite 230
; CITY: Fayetteville
; STATE: Arkansas AR
; COUNTRY: United States of America
; ZIP: 72701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS 6.2
; SOFTWARE: Wordperfect 6.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/191,593
; FILING DATE: 13 NOVEMBER 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/717,933
; FILING DATE: 23 SEPTEMBER 1996
; APPLICATION NUMBER: US 07/998,377
; FILING DATE: 30 DECEMBER 1992
; APPLICATION NUMBER: US 08/158,704
; FILING DATE: 29 NOVEMBER 1993
; APPLICATION NUMBER: US 60/009,455
; FILING DATE: 29 DECEMBER 1995
; APPLICATION NUMBER: US 08/610,424
; FILING DATE: 04 MARCH 1996
; ATTORNEY/AGENT INFORMATION:
```

```
NAME: ALEXANDER, DANIEL R
REGISTRATION NUMBER: 32,604
REFERENCE/DOCKET NUMBER: ARK00895601B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (501) 582-9111
TELEFAX: (501) 521-4931
TELEX: No. 6835824 applicable
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: No. 6835824 applicable
TOPOLOGY: unknown
MOLECULE TYPE: glycoprotein
DESCRIPTION: identified as derived N-terminal
DESCRIPTION: sequence of Ara h 2 p38.
HYPOTHETICAL: No
ANTI-SENSE: No. 6835824 applicable
FRAGMENT TYPE: N-terminal fragment
ORIGINAL SOURCE:
ORGANISM: Arachis hypogaea
STRAIN: Florunner
INDIVIDUAL ISOLATE: P38
DEVELOPMENTAL STAGE:
HAPLOTYPE: No. 6835824 applicable
TISSUE TYPE:
CELL TYPE: No. 6835824 applicable
CELL LINE: No. 6835824 applicable
ORGANELLE: No. 6835824 applicable
IMMEDIATE SOURCE:
LIBRARY:
CLONE: P38
POSITION IN GENOME:
CHROMOSOME/SEGMENT: No. 6835824 applicable
MAP POSITION: No. 6835824 applicable
UNITS: No. 6835824 applicable
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD: By agreement with
IDENTIFICATION METHOD: protein information and established
IDENTIFICATION METHOD: consensus sequence
OTHER INFORMATION: Seed storage protein and
OTHER INFORMATION: allergen
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-09-191-593-53
Alignment Scores:
Pred. No.: 3.38e-09 Length: 28
Score: 156.00 Matches: 28
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 12.2% Indels: 0
DB: 2 Gaps: 0
US-10-728-323-2 (1-717) x US-09-191-593-53 (1-28)
QY 59 CAGCAGTGGGAACCTCCAAAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTG 118
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 GlnGlnTrpGlnLeuGlnGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeu 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 119 AGGCCCTGCGAGCAACATCTCATG 142
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```



```
Db 21 ArgProCysGluGlnHisLeuMet 28
|||||
RESULT 12
US-08-670-186-4
; Sequence 4, Application US/08670186
; Patent No. 5859343
; GENERAL INFORMATION:
; APPLICANT: SUN, SAMUEL S.M.
; APPLICANT: XIONG, LIWEN
; APPLICANT: HU, ZHONG
; APPLICANT: CHEN, HANG
; TITLE OF INVENTION: RECOMBINANT SWEET PROTEIN MABINLIN
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 PENNSYLVANIA AVE NW, STE. 5500
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/670,186
; FILING DATE: 21-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 23461-20007.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 822-0168
; TELEX: 90-4030 MRSNFOERSWSH
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-670-186-4
Alignment Scores:
Pred. No.: 6,49e-09 Length: 158
Score: 156.00 Matches: 46
Percent Similarity: 47.6% Conservative: 32
Best Local Similarity: 28.0% Mismatches: 52
Query Match: 12.2% Indels: 34
DB: 1 Gaps: 8
US-10-728-323-2 (1-717) x US-08-670-186-4 (1-158)
QY 2 CTCACCATAGTAGCCCTCGCCCTTCCTCGCTGCCACGCATCTGCG---AGG 58
Db 5 IleLeuLeuLeuThrThrLeuAlaLeuPheValLeuAlaAsnAlaSerIleTyArg 24
QY 59 CAGCAGTGGGAACCTC-----CAAGGAGACAGAAAGATGCCAGAGC 97
Db 25 ThrThrValGluLeuAspGluGluAspAsnAspGluAsnGlnProLeuCysArg 44
QY 98 CAGCTCGAGAGG---CGGAACCTGAGGCCCTGGAGCAACATCTCATGCAAGATCCAA 154
Db 45 GlnPheGlnGlnHisGlnHisLeuArgAlaCysGlnArgTyIleArgArgAlaGln 64
QY 155 CGTGAC-----GAGGATTTCATATGAACGGGACCCGTCAGCCCTAGTCAGGATCCGTAC 208
Db 65 ArgGlyGlyValValAspGluLeuGluLeuGluAspValGluGluAsnGluAsp----- 82
QY 209 AGCCCTAGTCCATATGATCGGAGAGCGCTGGATCCTCTCAGCAACCAAGAGAGG----- 262
Db 83 -----GluAspGluAsnGlnGlnArgGlyPro 91
QY 263 -----TGTTGCAATGAGCTGAACGAGTTTGAGAACCAACAAAGGTGCATGTGC 310
Db 92 AlaLeuArgLeuCysCysAsnGlnLeuArgGlnVal-----AsnLysProCysValCys 109
QY 311 GAGGCATTGCAACAGATCATGAGAACCCAG-----AGCGATAGGTTCAGGGGAGGCAA 364
Db 110 ProValLeuArgGlnAlaAlaHisGlnGlnLeuTyArgGlnGlyGlnLeuGlyProArg 129
QY 365 CAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTGCCTCAACAGTGCAGCCTTAGGGCA 424
Db 130 GlnValArgGlnLeuPheArgAlaAlaArgAsnLeuProAsnIleCysLysIleProAla 149
QY 425 CCACAGCGTTGC 436
Db 150 ValGlyArgCys 153
RESULT 13
US-08-670-186-6
; Sequence 6, Application US/08670186
; Patent No. 5859343
; GENERAL INFORMATION:
; APPLICANT: SUN, SAMUEL S.M.
; APPLICANT: XIONG, LIWEN
; APPLICANT: HU, ZHONG
; APPLICANT: CHEN, HANG
; TITLE OF INVENTION: RECOMBINANT SWEET PROTEIN MABINLIN
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 PENNSYLVANIA AVE NW, STE. 5500
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/670,186
; FILING DATE: 21-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 23461-20007.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 822-0168
; TELEX: 90-4030 MRSNFOERSWSH
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-670-186-6
Alignment Scores:
Pred. No.: 2,38e-08 Length: 158
Score: 151.00 Matches: 45
Percent Similarity: 47.0% Conservative: 32
Best Local Similarity: 27.4% Mismatches: 53
Query Match: 11.8% Indels: 34
DB: 1 Gaps: 8
US-10-728-323-2 (1-717) x US-08-670-186-6 (1-158)
QY 2 CTCACCATAGTAGCCCTCGCCCTTCCTCGCTGCCACGCATCTGCG---AGG 58
```

Db	5	IleLeuLeuLeuThrLeuAlaLeuPheValIleuLeuAlaAsnAlaSerIleTyrArg	24
Qy	59	CAGCAGTGGGAATC-----CAAGGACAGAGAATGCCAGAGC	97
Db	25	ThrThrValGluLeuAspGluAspAsnAspGluAsnGlnProLeuCysArgArg	44
Qy	98	CAGCTCGAGAGG---GCGAACTGAGGCCCTGGCAGCAACATCTCATGCGAAGATCCAA	154
Db	45	GlnPheGlnGlnHisGlnGlnValArgAlaCysGlnArgTyrLeuArgArgAlaGln	64
Qy	155	CTGTGAC-----GAGATTTCATATGAACGGACCGGTACAGCCCTAGTCAGGATCCGTAC	208
Db	65	ArgGlyGlyLeuAlaAspGluLeuGluLeuAspValGluGluAsnGluAsp	82
Qy	209	AGCCCTAGTCCATATGATCGGAGAGGCCCTGGATCTCTCAGCACCAAGAGAGG---	262
Db	83	-----GluAspGluAsnGlnArgGlyPro	91
Qy	263	-----TGTTGCATGAGCTGAACAGATTTGAGAACCAACAAAGTGTCATGTGC	310
Db	92	AlaLeuArgLeuCysCysAsnGlnLeuArgGlnVal-----AsnLysProCysValCys	109
Qy	311	GAGGCATTCACACAGATCATGSGAGAACAG-----AGCGATAGTTGACGGGAGGCA	364
Db	110	ProValLeuArgGlnAlaAlaHisGlnGlnLeuTyrGlnGlyGlnIleGluGlyProArg	129
Qy	365	CAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCTCAACAGATCGCCCTAGGCGCA	424
Db	130	GlnValArgArgLeuPheArgAlaAlaArgAsnLeuProAsnIleCysLysIleProAla	149
Qy	425	CNACAGCTTGC	436
Db	150	ValGlyArgCys	153

RESULT 14

```

US-08-670-186-2
; Sequence 2, Application US/08670186
; Patent No. 585343
; GENERAL INFORMATION:
; APPLICANT: SUN, SAMUEL S.M.
; APPLICANT: XIONG, LIWEN
; APPLICANT: HU, ZHONG
; APPLICANT: CHEN, HANG
; TITLE OF INVENTION: RECOMBINANT SWEET PROTEIN MABINLIN
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 PENNSYLVANIA AVE NW, STE. 5500
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/670,186
; FILING DATE: 21-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 23461-20007.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 822-0168
; TELEX: 90-4030 MRSNFOERSH
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 155 amino acids

```

APPLICATION NUMBER: 07/742,834
FILING DATE: 08-AUGUST-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/550,804
FILING DATE: 09-JULY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/147,781
FILING DATE: 25-JANUARY-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/078,538
FILING DATE: 28-JULY-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 06/891,529
FILING DATE: 31-JULY-1986
ATTORNEY/AGENT INFORMATION:
NAME: Rae-Venter, Barbara
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: CGNE-037/04US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 926-6200
TELEFAX: (415) 854-3713
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 184 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-453-924-3

Alignment Scores:
Pred. No.: 5.7e-07 Length: 184
Score: 139.00 Matches: 45
Percent Similarity: 41.8% Conservative: 29
Best Local Similarity: 25.4% Mismatches: 57
Query Match: 10.9% Indels: 46
DB: Gaps: 8

US-10-728-323-2 (1-717) x US-08-453-924-3 (1-184)

QY	8	ATACTAGTACGCTCGCTTCCTCTCGCTGCCGACGATCTGCG---AGGCAGCAG	64
Db	8	ValSerAlaThrLeuAlaLeuPheLeuThrAsnAlaSerValTyrArgThrVal	27
QY	65	TGGGAACCTCAAGGAGAC-----AGAAAGATGC	91
Db	28	ValGluValAspGluAspAlaThrAsnProAlaGlyProPheArgIleProLysCys	47
QY	92	CAGAGCCAGCTCGAGGGCG---AACCTGAGCCCTGCGAGCAACATCTCATGCAGAAG	148
Db	48	ArgLysGluPheGlnGlnAlaGlnHisLeuLysAlaCysGlnGlnTrpLeuHisLysGln	67
QY	149	ATCCACAGTGAC-----GAGGATTCATATGAACGG	178
Db	68	AlaMetGlnSerGlySerGlyProSerTrpThrLeuAspGlyGluPheAspPheGluAsp	87
QY	179	GACCCGTACAGCCCTAGTCAGGATCGGTACAGCCCTAGTCCATATGATCGGAGGCGCT	238
Db	88	AspValGluAsnGlnGlnGlnGlyProGlnGlnArgProLeu-----	102
QY	239	GGATCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAACACCAA	298
Db	103	-----LeuGlnGlnCysCysAsnGluLeuHisGlnGlu-----Pro	115
QY	299	AGTGTCATGTGCGAGCATTTG-----CAACAGATCATGGAG	334
Db	116	LeuCysValCysProThrLeuLysGlyAlaSerLysAlaValLysGlnGlnIleArgGln	135
QY	335	AACCAAGCGATAGTTGCGAGGAGGCAACAGAGCAACAGTTCAAGAGG-----	385
Db	136	GlnGlnGlyGlnMetGlnGlnGlnMetGlnGlnMetGlnGlnValIleSerArgIleTyrGln	155
QY	386	GAGCTCAGGAACCTGCTCAACAGTGGCGCTTTAGGGCACACAGCGTTTC	436

Db 156 ThrAlaThrHisLeuProArgAlaCysAsnIleArgGlnValSerIleCys 172
Search completed: May 15, 2006, 22:28:14
Job time : 23.562 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: May 15, 2006, 23:34:07 ; Search time 20.2364 Seconds
(without alignments)
4441.255 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 1281
Sequence: 1 gctcaccatactagtagcccc.....taaaagatcatgttttgg 717

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues
Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DRV=xlp
-Q=/abss/ABSSWEB_spool/US10728323/runat_15052006_172145_22607/app_query.fasta_1
-DB=Published Applications_AA_Main -QFMT=fastan -SUFFIX=n2p.rapbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=biosum62 -TRANS=human4.0.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAESTZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss02p
-USER=US10728323 @CGN_1_1_503 @runat_15052006_172145_22607 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA_Main.*
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pcp.*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pcp.*
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pcp.*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pcp.*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	842	65.7	157	4	US-10-228-806-4
2	842	65.7	157	4	US-10-100-303A-63
3	838.5	65.5	207	5	US-10-358-324-2
4	835	65.2	160	5	US-10-899-551-4
5	819	63.9	156	4	US-10-245-871-53
6	819	63.9	156	4	US-10-253-286-53
7	771	60.2	166	4	US-10-100-303A-82
8	760	59.3	167	5	US-10-899-551-56
9	699	54.6	186	3	US-09-731-221-77
10	463	36.1	83	4	US-10-100-303A-81
11	243	19.0	158	4	US-10-302-633-2

12	243	19.0	158	4	US-10-424-599-169225	Sequence 169225,
13	241.5	18.9	155	4	US-10-424-599-264372	Sequence 264372,
14	177.5	13.9	168	4	US-10-165-289A-2	Sequence 2, Appli
15	149.5	11.7	167	4	US-10-382-066-29	Sequence 29, Appl
16	144	11.2	169	4	US-10-165-289A-4	Sequence 4, Appli
17	140.5	11.0	165	4	US-10-425-115-190971	Sequence 190971,
18	139	10.9	184	3	US-09-782-130-3	Sequence 3, Appli
19	139	10.9	184	4	US-10-100-121-48	Sequence 48, Appl
20	129.5	10.1	150	4	US-10-442-174A-1	Sequence 1, Appli
21	127.5	10.0	151	4	US-10-424-599-167336	Sequence 167336,
22	127	9.9	22	4	US-10-228-806-67	Sequence 67, Appl
23	126	9.8	102	4	US-10-424-599-219338	Sequence 219338,
24	124	9.7	148	4	US-10-425-115-219350	Sequence 219350,
25	119	9.3	20	3	US-09-731-221-66	Sequence 66, Appl
26	118	9.2	261	4	US-10-291-265-715	Sequence 715, App
27	117	9.1	20	3	US-09-731-221-63	Sequence 63, Appl
28	116	9.1	20	3	US-09-731-221-60	Sequence 60, Appl
29	115	9.0	20	3	US-09-731-221-61	Sequence 61, Appl
30	113	8.8	20	3	US-09-731-221-65	Sequence 65, Appl
31	112.5	8.8	101	4	US-10-424-599-275916	Sequence 275916,
32	112	8.7	20	3	US-09-731-221-54	Sequence 54, Appl
33	112	8.7	20	3	US-09-731-221-59	Sequence 59, Appl
34	112	8.7	20	3	US-09-731-221-62	Sequence 62, Appl
35	112	8.7	20	3	US-09-731-221-64	Sequence 64, Appl
36	111	8.7	20	3	US-09-731-221-57	Sequence 57, Appl
37	111	8.7	20	3	US-09-731-221-74	Sequence 74, Appl
38	111	8.7	20	3	US-09-731-221-75	Sequence 75, Appl
39	111	8.7	20	3	US-09-731-221-76	Sequence 76, Appl
40	110	8.6	25	4	US-10-245-871-60	Sequence 60, Appl
41	110	8.6	25	4	US-10-253-286-60	Sequence 60, Appl
42	109	8.5	20	4	US-10-245-871-55	Sequence 55, Appl
43	109	8.5	20	4	US-10-253-286-55	Sequence 55, Appl
44	109	8.5	149	5	US-10-481-032A-244	Sequence 244, App
45	108	8.4	20	3	US-09-731-221-53	Sequence 53, Appl

ALIGNMENTS

RESULT 1
US-10-228-806-4
; Sequence 4, Application US/10228806
; Publication No. US20030049237A1
; GENERAL INFORMATION:
; APPLICANT: Bannan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reactions
; FILE REFERENCE: 2002834-0043
; CURRENT APPLICATION NUMBER: US/10/228,806
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-228-806-4

Alignment Scores:
Pred. No.: 1.22e-79 Length: 157
Score: 842.00 Matches: 157
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 65.7% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-2 (1-717) x US-10-228-806-4 (1-157)

QY 2 CTCACCATACTAGTAGCCCTCGCCCTTTTCCTCGTCCGCCACGATCTGCGAGGCGAG 61
Db 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAATCTCAAGGACAGAGATGCCAGGACCTCGAGGGCGAAGCTTGAGG 121

Db 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCTCGGAGCAACATCTCATGCAAGAGATCCAAAGTGCAGGAGATTTCATATGAACGGGAC 181
Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 241
Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 80
QY 242 TCCTCTCAGCACCAAGAGAGGTGTTCGAATGAGCTGAACGAGTTTCAGAACCAACCAAGG 301
Db 81 SerSerGlnHisGlnGluArgCysAsnGlnLeuAsnGluPheGluAsnAsnGlnArg 100
QY 302 TGCATGTGCGAGGCATTGCAACAGATCATGGAAGAACACGAGCGGATAGTTGCAGGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGAGCAACAGTTCAAGAGGAGCTCAGGAACCTGCCTCAACAGTCCGCGCCTTAGG 421
Db 121 GlnGlnGluGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACCAAGAGGTTGCAGCTTGGACGTCGAAAGTGGCGGAGAGACAGATAC 472
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157
RESULT 2
US-10-100-303A-63
; Sequence 63, Application US/10100303A
; Publication No. US20030202980A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; FILE REFERENCE: to Allergy
; CURRENT APPLICATION NUMBER: US/10/100,303A
; CURRENT FILING DATE: 2002-03-18
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Arachis hypogaea, Prot/Nucleo Ara h 2
US-10-100-303A-63
Alignment Scores:
Pred. No.: 1.22e-79 Length: 157
Score: 842.00 Matches: 157
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 65.7% Indels: 0
DB: 4 Gaps: 0
US-10-728-323-2 (1-717) x US-10-100-303A-63 (1-157)
QY 2 CTCACCATACTAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCATCTGCGAGGAG 61
Db 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAATCCCAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTTGAGG 121
Db 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCTCGGAGCAACATCTCATGCAAGAGATCCAAAGTGCAGGAGATTTCATATGAACGGGAC 181
Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGGGCTGGA 241
Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 80
QY 242 TCCTCTCAGCACCAAGAGAGGTGTTCGAATGAGCTGAACGAGTTTCAGAACCAACCAAGG 301

Db 81 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100
QY 302 TGCATGTGCGAGGCATTGCAACAGATCATGGAAGAACACGAGCGGATAGTTGCAGGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTGCCTCAACAGTCCGCGCCTTAGG 421
Db 121 GlnGlnGluGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACCAAGAGGTTGCAGCTTGGACGTCGAAAGTGGCGGAGAGACAGATAC 472
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157
RESULT 3
US-10-958-324-2
; Sequence 2, Application US/10958324
; Publication No. US20050114924A1
; GENERAL INFORMATION:
; APPLICANT: DODO, HORTENSE W.
; APPLICANT: ARNTZEN, CHARLES J.
; APPLICANT: KONAN, KOFFI N'DA
; APPLICANT: VIOUEZ, OLGA
; TITLE OF INVENTION: DOWN-REGULATION AND SILENCING OF ALLERGEN GENES IN
; FILE REFERENCE: 072121/0104
; CURRENT APPLICATION NUMBER: US/10/958,324
; CURRENT FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US/09/715,036
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/167,255
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-958-324-2
Alignment Scores:
Pred. No.: 3.08e-79 Length: 207
Score: 838.50 Matches: 162
Percent Similarity: 90.3% Conservatives: 5
Best Local Similarity: 87.6% Mismatches: 17
Query Match: 65.5% Indels: 1
DB: 5 Gaps: 1
US-10-728-323-2 (1-717) x US-10-958-324-2 (1-207)
QY 2 CTCACCATACTAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCATCTGCGAGGAG 61
Db 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 23
QY 62 CAGTGGGAATCCCAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTTGAGG 121
Db 24 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
QY 122 CCTCGGAGCAACATCTCATGCAAGAGATCCAAAGTGCAGGAGATTTCATATGAACGGGAC 181
Db 44 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 63
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGGGCTGGA 241
Db 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 83
QY 242 TCCTCTCAGCACCAAGAGAGGTGTTCGAATGAGCTGAACGAGTTTCAGAACCAACCAAGG 301
Db 84 SerSerGlnHisGlnGluArgCysAsnGlnLeuAsnGluPheGluAsnAsnGlnArg 103
QY 302 TGCATGTGCGAGGCATTGCAACAGATCATGGAAGAACACGAGCGGATAGTTGCAGGGAGG 361
Db 104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123

[illegible]

```

RESULT 4
US-10-899-551-4
; Sequence 4, Application US/10899551
; Publication No. US20050063994A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Burks, A. Wesley
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Howard, Sosin B.
; APPLICANT: Bottomly, Kim H.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy
; FILE REFERENCE: 2002834-0233
; CURRENT APPLICATION NUMBER: US/10/899,551
; CURRENT FILING DATE: 2004-07-26
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 160
; TYPE: PRT
; ORGANISM: species Arachis hypogaea
US-10-899-551-4

```

Alignment Scores:	
Pred. No.:	6,73e-79
Score:	835.00
Length:	160
Percent Similarity:	100.0%
Matches:	156
Best Local Similarity:	100.0%
Conservative:	0
Mismatches:	0
Indels:	0
Query Match:	65.2%
Gaps:	0
DB:	5

US-10-728-323-2 (1-717) x US-10-899-551-4 (1-160)

Qy	2	CTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCTCGTCGCCACGCATCTCGAGGCAG	61
Db	4	LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaAraGln	23
Qy	62	CAGTGGGAACCTCAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAAACCTGAGG	121
Db	24	GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluAraGlaAsnLeuAraG	43
Qy	122	CCCTCGCGACCACTCTATCGAGAGGATCCAACTGACGAGGATTCATATGAACCGGAC	181
Db	44	ProCysGluGlnHisLeuMetGlnLysIleGlnAraGspGluAspSerTyrGluAraGsp	63
Qy	182	CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGCGCTGGA	241
Db	64	ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly	83
Qy	242	TCCTCTCAGCACCAAGAGAGGTGTTCGAATCAGCTGAACGAGTTTGAACAACCAAGG	301
Db	84	SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnAraG	103
Qy	302	TGCATGTGGCGGCATTGCCACAGATCATGAGAACCGAGCCGATAGTTTCGAGGGGAGG	361
Db	104	CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyAraG	123

Qy	362	CAACAGAGCAACAGTTCAAGAGGAGCTCAGGAACCTGCCTCAACAGTGGCGCCTTAGG	421
Db	124	GlnGlnGluGlnGlnPheLysArgGluLeuArgAenLeuProGlnGlnCysGlyLeuArg	143
Qy	422	GCACCCAGAGCGTTGCGAGCTTGGAGCTCGAAAGTGCAGCGCAGACAGACA	469
Db	144	AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArg	159

RESULT 5
US-10-245-871-53
; Sequence 53, Application US/10245871
; Publication No. US20030235594A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: 11-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2013
; CURRENT APPLICATION NUMBER: US/10/245,871
; CURRENT FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-245-871-53

Alignment Scores:		
Pred. No.:	3,26e-77	Length: 156
Score:	819.00	Matches: 153
Percent Similarity:	100.0%	Conservative: 0
Best Local Similarity:	100.0%	Mismatches: 0
Query Match:	63.9%	Indels: 0
DB:	4	Gaps: 0

US-10-728-323-2 (1-717) x US-10-245-871-53 (1-156)

Qy	2	CTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCTCGTGCACGCATCTCGAGGCAG	61
Db	4	LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaargGln	23
Qy	62	CAGTGGAACTCCAAGGAGCAGAGAAGATGCCAGAGCCAGCTCGAGAGGCGCAACTGTAGG	121
Db	24	GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuAArg	43
Qy	122	CCCTGCGAGCAACATCTCATGCAGAAGATCCAACGTGACGAGGATTTCATATCAACGGGAC	181
Db	44	ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp	63
Qy	182	CCGTACAGCCCTAGTCAAGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCCCTCGA	241
Db	64	ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly	83
Qy	242	TCCTCTCAGCACCAAGAGAGGTGTTCGAATGAGCTGAAACGAGTTTGAGAACCAACAAAGG	301
Db	84	SerSerGlnHisGlnGluAArgCysCysAsnGluLeuAsnGluPheGluAsnAenGlnAArg	103
Qy	302	TGCATGTCCGAGGCAATTGCCAACAGATCATGAGAACCCAGAGCGGATAGTTGACAGGGGAGG	361
Db	104	CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyAArg	123
Qy	362	CAACAGGAGCAACAGTTTCAAGAGGGAGCTCAGGAACTTGCCTCAACAGTGCGGCCCTTAGG	421
Db	124	GlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuAArg	143
Qy	422	GCACCAGCGTTCCGACTTCGACGTTCGAAAGTGGCGGC	460
Db	144	AlaProGlnArgCysAspLeuAspValGluSerGlyGly	156

US-10-899-551-56

Alignment Scores:
Pred. No.: 5 58e-71 Length: 167
Score: 760.00 Matches: 139
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 59.3% Indels: 0
DB: 5 Gaps: 0

US-10-728-323-2 (1-717) x US-10-899-551-56 (1-167)

QY 53 GCGAGGAGGAGTGGGAATCTCAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGGCG 112
Db 17 AlaArgGlnGlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAla 36
QY 113 AACCTGAGCCCTGCGAGCAACATCTCATGCAGAGATCCACGTCACGAGGATTCATAT 172
Db 37 AsnLeuArgProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyr 56
QY 173 GAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATATGATCGAGA 232
Db 57 GluArgAspProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArg 76
QY 233 GCGCTGGATCTCTCAGCACCAAGAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAAC 292
Db 77 GlyAlaGlySerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsn 96
QY 293 AACCAAGGTGCTGTGCGAGGATTCGACAGATTCGACAGATCATGAGAACCCAGAGCGATAGTTG 352
Db 97 AsnGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeu 116
QY 353 CAGGGGAGGCAACAGAGGCAACAGTTCAAGAGGGAGCTCAGGAACCTGCCTCAACAGTGC 412
Db 117 GlnGlyArgGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCys 136
QY 413 GCGCTTAGGCAACACAGAGCTGCGACTTGGAGCTCGAAAGTGGCGGCAGAGACAGA 469
Db 137 GlyLeuArgAlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArg 155

RESULT 9

US-09-731-221-77
; Sequence 77, Application US/09731221
; Patent No. US20020018778A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael
; TITLE OF INVENTION: Passive Desensitization
; FILE REFERENCE: 2002834-0103
; CURRENT APPLICATION NUMBER: US/09/731,221
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 77
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Arachis
; OTHER INFORMATION: Hypogaea
US-09-731-221-77

Alignment Scores:
Pred. No.: 1 52e-64 Length: 166
Score: 699.00 Matches: 131
Percent Similarity: 92.9% Conservative: 0
Best Local Similarity: 92.9% Mismatches: 10
Query Match: 54.6% Indels: 0
DB: 3 Gaps: 0

US-10-728-323-2 (1-717) x US-09-731-221-77 (1-166)

QY 50 TCTGCGAGGAGGAGTGGGAATCTCAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGG 109
|||||

Db 15 SerAlaArgGlnGlnAlaGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuAlaArg 34
QY 110 GCGAACCTGAGGCCCTGCGAGCAACATCTCATGCAGAAAGATCCAAACGTCAGCAGGATTC 169
Db 35 AlaAsnLeuArgAlaCysGluAlaHisLeuMetGlnLysIleGlnAlaAspGluAspSer 54
QY 170 TATGAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATATGATCG 229
Db 55 TyrGluArgAlaProTyrSerProSerGlnAlaProTyrSerProSerProTyrAspArg 74
QY 230 AGAGCGGCTGGATCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGACCAAGCTTGG 289
Db 75 ArgGlyAlaGlySerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGlu 94
QY 290 AACCAACAAAGGTGATGTGCGAGGATTCGAAACAGATCATGCAGAACCCAGAGCGATAGG 349
Db 95 AsnAsnGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArg 114
QY 350 TTGCAAGGAGGCAACAGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTGCCTCAACAG 409
Db 115 LeuGlnGlyAlaGlnGlnGlnPheLysArgGluAlaArgAsnLeuProGlnGln 134
QY 410 TCGGCGCTTAGGCAACACAGAGCGTTGCGACTTGGAGCTCGAAAGTGGCGGCAGAGACAGA 469
Db 135 CysGlyLeuArgAlaProGlnArgCysAspAlaAspValGluSerGlyGlyArgAspArg 154
QY 470 TAC 472
|||
Db 155 Tyr 155

RESULT 10

US-10-100-303A-81
; Sequence 81, Application US/10100303A
; Publication No. US20030202980A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; FILE REFERENCE: to Allergy
; FILE REFERENCE: 2002834-0166
; CURRENT APPLICATION NUMBER: US/10/100,303A
; CURRENT FILING DATE: 2002-03-18
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Arachis hypogaea, Prot/Nucleo Ara h 2
US-10-100-303A-81

Alignment Scores:
Pred. No.: 1e-39 Length: 83
Score: 463.00 Matches: 83
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 36.1% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-2 (1-717) x US-10-100-303A-81 (1-83)

QY 68 GAATCTCAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGCGCAACCTGAGGCCCTGC 127
Db 1 GluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArgProCys 20
QY 128 GAGCAACATCTCATGCAGAAAGATCCAAACGTCAGCAGGATTCATATGAACGGGACCGGTAC 187
Db 21 GluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAspProTyr 40
QY 188 AGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGGGCGCTGCATCCTCT 247
Db 41 SerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGlySerSer 60
QY 248 CAGCAACCAAGAGAGGTGTGCAATGAGCTGAACCGAGTTTGAGAAACCAACCAAGAGTGCATG 307
|||||

Db 61 GlnHisGlnArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArgCysMet 80

QY 308 TGGAGGCA 316
|||||

Db 81 CysGluAla 83

RESULT 11

US-10-302-633-2

; Sequence 2, Application US/10302633

; Publication No. US20030229038A1

; GENERAL INFORMATION:

; APPLICANT: de Lumen, Benito O.

; TITLE OF INVENTION: Lunasin Peptides

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 75 DENISE DRIVE

; CITY: HILLSBOROUGH

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94010

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/302,633

; FILING DATE: 22-Mar-2000

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/531,727

; FILING DATE: 21-Mar-2000

; APPLICATION NUMBER: 08/938,675

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: OSMAN, RICHARD A

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: B98-003

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (650) 343-4341

; TELEFAX: (650) 343-4342

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 158 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-302-633-2

Alignment Scores:

Pred. No.: 1.94e-16 Length: 158

Score: 243.00 Matches: 56

Percent Similarity: 56.5% Conservative: 31

Best Local Similarity: 36.4% Mismatches: 55

Query Match: 19.0% Indels: 12

DB: 4 Gaps: 6

US-10-728-323-2 (1-717) x US-10-302-633-2 (1-158)

QY 5 ACCATACTAGTACGCGCTTTCCTCTCGCTGCCACGATCTGGAGGACGAG 64

Db 5 ThrileuLeuIleSerLeuPheCysile---AlaHisThrCysSerAlaSerLys 23

QY 65 TGGAACTCCAAGAGACAGAGATGCCAGAGCCAGCTCGAGGGCGAACCCTGAGGCC 124

Db 24 TrpGlnHisGlnAsp---SerCysArgLysGlnLeuGlnGlyValAsnLeuThrPro 42

QY 125 TGGCAGCAACATCTCATGCAAGATCCAA-----CGTGACGAGGATTCATATGAACGG 178

Db 43 CysGluLysHisIleMetGluLysIleGlnGlyArgLysAspAspAspAsp 62

QY 179 GACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGGCGCT 238

Db 63 AspAspAsnHisIleLeuArgThrMetArgGlyArgIleAsnTyrlleArgAsnGlu 82

QY 239 GGATCTCTCTCAG-----CACCAGAGAGGCTTTCGCAATGAGCTGAAC 280

Db 43 CysGluLysHisIleMetGluLysIleGlnGlyArgGlyAspAspAspAspAsp 62

QY 179 GACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGGCGCT 238

Db 63 AspAspAsnHisIleLeuArgThrMetArgGlyArgIleAsnTyrlleArgAsnGlu 82

QY 239 GGATCTCTCTCAG-----CACCAGAGAGGCTTTCGCAATGAGCTGAAC 280

Db 83 GlyLysAspGluAspGluGluGluGlyHisMetGlnLysCysCysThrGluMetSer 102

QY 281 GAGTTTGTGAGAACCAACAAAGTTCATGTCGAGGATTCACACAGATCATGGAGAACCCAG 340

Db 103 GluLeu---ArgSerProLysCysGlnCysLysAlaLeuGlnLysIleMetGluAsnGln 121

QY 341 AGCGATAGTTTCAGGGGAGGCAACAGAGGACCAACAGTTCAGAGGAGGCTCAGGAACCTTG 400

Db 122 SerGluGluLeuGlu---GluLysGlnLysLysLysMetGluLysGluLeuIleAsnLeu 140

QY 401 CCTCAACAGTGGCGCTTAGGCGCACACAGCGTTGCGACTTG 442

Db 141 AlaThrMetCysArgPheGlyProMetIleGlnCysAspLeu 154

RESULT 12

US-10-424-599-169225

; Sequence 169225, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 169225

; LENGTH: 158

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_123825C.1.pap

US-10-424-599-169225

Alignment Scores:

Pred. No.: 1.94e-16 Length: 158

Score: 243.00 Matches: 56

Percent Similarity: 56.5% Conservative: 31

Best Local Similarity: 36.4% Mismatches: 55

Query Match: 19.0% Indels: 12

DB: 4 Gaps: 6

US-10-728-323-2 (1-717) x US-10-424-599-169225 (1-158)

QY 5 ACCATACTAGTACGCGCTTTCCTCTCGCTGCCACGATCTGGAGGACGAG 64

Db 5 ThrileuLeuIleSerLeuPheCysile---AlaHisThrCysSerAlaSerLys 23

QY 65 TGGAACTCCAAGAGACAGAGATGCCAGAGCCAGCTCGAGGGCGAACCCTGAGGCC 124

Db 24 TrpGlnHisGlnAsp---SerCysArgLysGlnLeuGlnGlyValAsnLeuThrPro 42

QY 125 TGGCAGCAACATCTCATGCAAGATCCAA-----CGTGACGAGGATTCATATGAACGG 178

Db 43 CysGluLysHisIleMetGluLysIleGlnGlyArgLysAspAspAspAsp 62

QY 179 GACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGGCGCT 238

Db 63 AspAspAsnHisIleLeuArgThrMetArgGlyArgIleAsnTyrlleArgAsnGlu 82

QY 239 GGATCTCTCTCAG-----CACCAGAGAGGCTTTCGCAATGAGCTGAAC 280

```

Db      83  GlyIysAspGluuAspGluuGluGluGluYHisMetGlnLysCysCysThrGluMetSer 102
Qy      281  GAGTTTCAGAAACACCAAGAGTGTCATGTGCGAGGCATTGCAACAGATCATGTGAGAACCCAG 340
Db      103  GluLeu--ArgSerProLysCysGlnCysLysAlaLeuGlnLysIleMetGluAsnGln 121
Qy      341  AGCATAGGTTGCAGGGGAGGCACAGGACACAGTTCAAGAGGGAGCTCAGGAACCTTG 400
Db      122  SerGluLeuLeuGlu--GluLysGlnLysLysMetGluLysGluLeuIleAsnLeu 140
Qy      401  CCTCAACAGTGCGGCCCTTAGCGGCACACACAGCGTTCGCACTTG 442
Db      141  AlaThrMetCysArgPheGlyProMetIleGlnCysAspLeu 154

RESULT 13
US-10-424-599-264372
; Sequence 264372, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 264372
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; ; OTHER INFORMATION: Clone ID: PAT_MRT3847_80749C.1.pep
US-10-424-599-264372

```

Alignment Scores:					
Pred. No.:	2,77e-16	Length:	155		
Score:	241.50	Matches:	58		
Percent Similarity:	57.1%	Conservative:	30		
Best Local Similarity:	37.7%	Mismatches:	53		
Query Match:	18.9%	Indels:	13		
DB:	4	Gaps:	7		
 US-10-728-323-2 (1-717) x US-10-424-599-264372 (1-155)					
Qy	2	CTCACCATAGTAGTCCGTCGCCCTTTTCTCTCCTCGTGCCCACGCATCTCGAGGCAG	61		
Db	4	LeuthrileLeuLeuAlaLeuPhele-----AlaHisThrcysCysalaSer	21		
Qy	62	CAGTGGGAACCTCAAGGAGACAGAAGATGCCAGACCAGCTCGAGAGGCGCAACTGTAGG	121		
Db	22	LysTrpGlnGlnHisIsgLnGlnGluSerCysArgGluGlnLeuLysGlyleAAsnLeuAen	41		
Qy	122	CCCTCGAGCAACATCTCATGCAGAAGATCCAA-----CGTCAGCAGGATTCATAT	172		
Db	42	ProCysGlu---HisIleMetGlnLysIleGlnAlaGlyArgArgGlyGlnAspGlySer	60		
Qy	173	GAA CGG G A C C C G T C A C A G C C T A G T C A G G A T C C C T A G C C C T A G T C C C A T G A T C G G A G A	232		
Db	61	AspGluAspHisIleLeuIleArgThrMetPro---GlyArgIleAsnTyrlleArgLys	79		
Qy	233	GGCGCTGATCCTCTCAG-----CACCAGAGAGAGTGTTCNAATCAGCTGAAC	280		
Db	80	LysGluGlyLysGlnGluGluGluGluGlyHisMetGlnLysCysCysSerGluMetSer	99		
Qy	281	GAGTTTGTAGACAACAACAAAGGTGCATTGTGCGAGGCATTGCAACAGATCATGAGAACACAG	340		
Db	100	GluLeuLys---SerProIleCysGlnCysLysAlaLeuGlnLysIleMetAspIasnGln	118		
Qy	341	AGCGATAGTGTTCAGGGGGGCAACAGGAGCAACAGTTTCAGAGGGAGGCTCAGGAAC TTG	400		

```

Db      119 SerGluGlnLeuGluGlyLys--GluLysLysGlnMetGluArgGluLeuMetAsnLeu 137
Qy      401 CCTCAACAGTCGCGCCTTAGGCGACCACACAGCGTTTCGCACTTG 442
          ::||| ||| |||||
Db      138 AlaIleArgCysArgLeuGlyProMetIleGlyCysAspLeu 151

RESULT 14
US-10-165-289A-2
; Sequence 2, Application US/10165289A
; Publication No. US20030159174A1
; GENERAL INFORMATION:
;   APPLICANT: QIU, Xiao
;   APPLICANT: TRUSKA, Martin
;   APPLICANT: HU, Zhiyuan
;   TITLE OF INVENTION: Flax (Linum usitatissimum L.)
;   FILE OF INVENTION: Seed-Specific Promoters
;   FILE REFERENCE: BNZ-005
;   CURRENT APPLICATION NUMBER: US/10/165,289A
;   CURRENT FILING DATE: 2002-06-06
;   PRIOR APPLICATION NUMBER: 60/295823
;   PRIOR FILING DATE: 2001-06-06
;   NUMBER OF SEQ ID NOS: 13
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO 2
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Linum usitatissimum
US-10-165-289A-2

Alignment Scores:
Pred No.: 1,6e-09      Length: 168
Score: 177,50      Matches: 54
Percent Similarity: 45.7%      Conservative: 26
Best Local Similarity: 30.9%      Mismatches: 48
Query Match: 13.9%      Indels: 47
DB: 4      Gaps: 9

```

Db 119 SerGluGlnLeuGluGlyLys--GluLysLysGlnMetGluArgGluLeuMetAsnLeu 13
 Qy 401 CCTCAACAGTCGGCCTTATAGGCACACAGCGTTTCGCACTTG 442
 Db 138 AlaIleArgCysArgLeuGlyProMetIleGlyCysAspLeu 151
 RESULT 14
 US-10-165-289A-2
 ; Sequence 2, Application US/10165289A
 ; Publication No. US20030159174A1
 ; GENERAL INFORMATION:
 ; APPLICANT: QIU, Xiao
 ; APPLICANT: TRUSKA, Martin
 ; APPLICANT: HU, Zhiyuan
 ; TITLE OF INVENTION: Flax (Linum usitatissimum L.)
 ; TITLE OF INVENTION: Seed-Specific Promoters
 ; FILE REFERENCE: BNZ-005
 ; CURRENT APPLICATION NUMBER: US/10/165,289A
 ; CURRENT FILING DATE: 2002-06-06
 ; PRIOR APPLICATION NUMBER: 60/295823
 ; PRIOR FILING DATE: 2001-06-06
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 168
 ; TYPE: PRT
 ; ORGANISM: Linum usitatissimum
 US-10-165-289A-2
 Alignment Scores:
 Pred. No.: 1.6e-09 Length: 168
 Score: 177.50 Matches: 54
 Percent Similarity: 45.7% Conservative: 26
 Best Local Similarity: 30.9% Mismatches: 48
 Query Match: 13.9% Indels: 47
 DB: 4 Gaps: 9
 US-10-728-323-2 (1-717) x US-10-165-289A-2 (1-168)
 Qy 2 CTCACCATACTAGTAGCCCTCGCC-----CTTTTCCTCCTCGCTGCCCAACGCATCT 52
 Db 4 LeuMetSerLeuAlaAlaValAlaThrAlaPheLeuPheLeuIleValValAspIleSer 23
 Qy 53 GCGAGCGCAGTAGTGGGAATCCAAAGGAGAC----- 82
 Db 24 ValArgThrValIleAlaAspGluAspThrAsnGlnGlyArgGlyGlnGlyGly 43
 Qy 83 -----AGAAAGATGCCAGAGCCAGCTC---GAGAGGCGNAACCTGAGGCCCTGC 127
 Db 44 GlnGlyGlnGlnGlnGlnCysGluLysGlnIleGlnGluGlnAspTyrLeuArgSerCys 63
 Qy 128 GAGCAACATCTCATGCAGAAATCCAACTGACGAGGATTTCATATCAACGGGACCCGCTAC 187
 Db 64 GlnGlnPheLeuTrpGluLysValGlnLysGlyLysGlyArgSerTyr----- 78
 Qy 188 AGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGGCCCTCGA----- 241
 Db 79 -----TyrTyrAsnGlnGlyArgGlyGlyGly 87
 Qy 242 ---TCCTCTCAGCACCAGAGAGGTGTTGCAATGAGCTGAACAGTTCAGAACACCAA 298
 Db 88 GlnGlnSerGlnHisPheAspSerCysAspLeuLysGlnLeuArgSer----- 105
 Qy 299 AGGTGCATGTCGGAGGCATTCGAACAGATCATG-----GACAACCAAGAGCGATAGTTG 352
 Db 106 GluCysThrCysArgGlyLeuGluArgAlaIleGlyGlnMetArgGlnAspIleGlnGln 125
 Qy 353 CAGGGGAGGCAACAGGAG-----CAACAGTTCAAGAGGGAGCTCAGGAAC 397
 Db 126 GlnGlyGlnGlnGlnGluValGluArgTrpValGlnGlnAlaLysGlnValAlaArgAsp 145
 Qy 398 TTGCCTCAACAGTGGCGCCTTAGGCAACACAGCGTTGCGACTTG 442
 Db 119 SerGluGlnLeuGluGlyLys--GluLysLysGlnMetGluArgGluLeuMetAsnLeu 13
 Qy 401 CCTCAACAGTCGGCCTTATAGGCACACAGCGTTTCGCACTTG 442
 Db 138 AlaIleArgCysArgLeuGlyProMetIleGlyCysAspLeu 151

Search completed: May 16, 2006, 00:00:30
Job time : 104.182 secs

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 23:41:32 ; Search time 2.46663 Seconds
(without alignments)
4094.130 Million cell updates/sec

Title: US-10-728-323-2

Perfect score: 1281

Sequence: 1 gctcaccatactagtagccccc.....taaaagatcatgtttgtttgt 717

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapcp 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 250354 seqs, 46948837 residues

Total number of hits satisfying chosen parameters: 500708

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-O=/abs/ABSWEB spool/US10728323/runat_15052006_172148_22650/app_query.fasta.1
-DB=Published Applications_AA_New -GPM=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abs02p
-USER=US10728323 @CGN 1.1.63 @runat_15052006_172148_22650 -NCPU=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database : Published Applications_AA_New.*

1: /SID55/ptodata/1/pubpaa/US08_NEW_PUB.pep1.*
2: /SID55/ptodata/1/pubpaa/US06_NEW_PUB.pep1.*
3: /SID55/ptodata/1/pubpaa/US07_NEW_PUB.pep1.*
4: /SID55/ptodata/1/pubpaa/US08_NEW_PUB.pep1.*
5: /SID55/ptodata/1/pubpaa/PCT_NEW_PUB.pep1.*
6: /SID55/ptodata/1/pubpaa/US09_NEW_PUB.pep1.*
7: /SID55/ptodata/1/pubpaa/US09_NEW_PUB.pep1.*
8: /SID55/ptodata/1/pubpaa/US10_NEW_PUB.pep1.*
9: /SID55/ptodata/1/pubpaa/US10_NEW_PUB.pep1.*
10: /SID55/ptodata/1/pubpaa/US11_NEW_PUB.pep1.*
11: /SID55/ptodata/1/pubpaa/US11_NEW_PUB.pep1.*
12: /SID55/ptodata/1/pubpaa/US60_NEW_PUB.pep1.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	819	63.9	156	11	US-11-033-039-53
2	243	19.0	158	9	US-10-508-263-44
3	241.5	18.9	155	9	US-10-508-263-42
4	184.5	14.4	323	9	US-10-508-263-52
5	151	11.8	295	9	US-10-508-263-50

6	146.5	11.4	164	9	US-10-508-263-2	Sequence 2, Appli
7	118	9.2	261	11	US-11-000-463-715	Sequence 715, App
8	116.5	9.1	319	11	US-11-096-568A-24535	Sequence 24535, A
9	110	8.6	25	11	US-11-033-039-60	Sequence 60, Appl
10	109	8.5	20	11	US-11-033-039-55	Sequence 55, Appl
11	107.5	8.4	260	11	US-11-030-308-7	Sequence 7, Appli
12	106.5	8.3	178	9	US-10-508-263-32	Sequence 32, Appli
13	105.5	8.2	164	9	US-10-508-263-4	Sequence 4, Appli
14	105.5	8.2	180	9	US-10-508-263-38	Sequence 38, Appl
15	104	8.1	144	9	US-10-508-263-40	Sequence 40, Appl
16	103.5	8.1	166	9	US-10-508-263-8	Sequence 8, Appli
17	103.5	8.1	178	9	US-10-508-263-34	Sequence 34, Appli
18	102.5	8.0	178	9	US-10-508-263-46	Sequence 46, Appl
19	101.5	7.9	178	9	US-10-508-263-36	Sequence 36, Appl
20	99.5	7.8	148	9	US-10-526-716-2	Sequence 2, Appli
21	96	7.5	144	9	US-10-508-263-48	Sequence 48, Appl
22	95	7.4	170	9	US-10-508-263-6	Sequence 6, Appli
23	95	7.4	568	9	US-10-784-004-390	Sequence 390, App
24	92	7.2	336	11	US-11-172-740-126	Sequence 126, App
25	91	7.1	297	9	US-10-203-486-3	Sequence 3, Appli
26	90	7.0	448	11	US-11-119-569-21	Sequence 21, Appl
27	90	7.0	448	11	US-11-119-569-4	Sequence 4, Appli
28	89	6.9	1341	11	US-11-188-298-21361	Sequence 21361, A
29	87.5	6.8	368	11	US-11-188-298-13415	Sequence 13415, A
30	87	6.8	585	9	US-10-784-004-712	Sequence 712, App
31	87	6.8	1236	9	US-10-873-528-109	Sequence 109, App
32	86.5	6.8	628	11	US-11-183-136-10	Sequence 10, Appl
33	86	6.7	1230	11	US-11-087-099-1702	Sequence 1702, Ap
34	86	6.7	1230	11	US-11-087-099-8922	Sequence 8922, Ap
35	86	6.7	1230	11	US-11-188-298-1685	Sequence 1685, Ap
36	86	6.7	1230	11	US-11-188-298-8275	Sequence 8275, Ap
37	85.5	6.7	185	11	US-11-264-096-1711	Sequence 1711, Ap
38	85.5	6.7	378	11	US-11-172-740-1335	Sequence 1335, Ap
39	85.5	6.7	638	11	US-11-072-512-2626	Sequence 2626, Ap
40	85.5	6.7	1162	10	US-11-106-014-94	Sequence 94, Appl
41	85	6.6	15	11	US-11-033-039-54	Sequence 54, Appl
42	85	6.6	20	11	US-11-033-039-59	Sequence 59, Appl
43	84.5	6.6	280	11	US-11-072-512-3063	Sequence 3063, Ap
44	84.5	6.6	716	11	US-11-194-890-20	Sequence 20, Appl
45	83.5	6.5	289	11	US-11-096-568A-8938	Sequence 8938, Ap

ALIGNMENTS

RESULT 1
US-11-033-039-53
; Sequence 53, Application US/11033039
; Publication No. US20060002947A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: LI-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2017US01
; CURRENT APPLICATION NUMBER: US/11/033,039,
; CURRENT FILING DATE: 2005-01-11
; PRIOR APPLICATION NUMBER: 10/245,871
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 1452
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 53
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-11-033-039-53

Alignment Scores:
Pred. No.: 2,018-73 Length: 156
Score: 819.00 Matches: 153
Percent Similarity: 100.0% Conservative: 0

```

Best Local Similarity: 100.0%      Mismatches: 0
Query Match: 63.9%               Indels: 0
DB: 11                          Gaps: 0

US-10-728-323-2 (1-717) x US-11-033-039-53 (1-156)

QY 2 CTCACCATAGTAGCCCTCGCCCTTTCCTCTCGCTGCCACGCATCTGGAGGAG 61
Db 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisSerAlaArgGln 23
QY 62 CAGTGGAACTCAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTCAGG 121
Db 24 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
QY 122 CCTCGGAGCAACATCTTCATGCAAGATCCACGTGACGAGATTCATATGAACGGGAC 181
Db 44 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 63
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGGCGCTGGA 241
Db 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrSerProTyrSerProTyr 83
QY 242 TCCTCTCAGACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAACCAACAAAGG 301
Db 84 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 103
QY 302 TGCATGTGCGAGCATTTGCAACAGATCATGGAAACAGAGCCGATAGGTTGCGAGGGAGG 361
Db 104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123
QY 362 CACAGAGCAACAGTTCAGAGGAGCTCAGAACTTGCTCAACAGTCCGCGCTTAGG 421
Db 124 GlnGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143
QY 422 GCACCAAGGCTTGGCACTTGGACGTGCGAAGTGGCGGC 460
Db 144 AlaProGlnArgCysAspLeuAspValGluSerGlyGly 156

RESULT 2
US-10-508-263-44
; Sequence 44, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Glycine max
US-10-508-263-44

Alignment Scores:
Pred. No.: 4,06e-16      Length: 158
Score: 243.00           Matches: 56
Percent Similarity: 56.5%      Conservative: 31
Best Local Similarity: 36.4%   Mismatches: 55
Query Match: 19.0%          Indels: 12
DB: 9                   Gaps: 6

US-10-728-323-2 (1-717) x US-10-508-263-44 (1-158)

QY 5 ACCATAGTAGCCCTCGCCCTTTCCTCTCGCTGCCACGCATCTGGAGGAGCAG 64
Db 5 ThrIleLeuLeuSerLeuLeuPheCysIle---AlaHisThrCysSerAlaSerLys 23
QY 65 TGGGAACCTCAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGGCGAACCTGAGGCC 124
Db 24 TrpGlnHisGlnGlnAsp---SerCysArgLysGlnLeuGlnGlyValAsnLeuThrPro 42

```

```

QY 125 TCGAGCAACATCTCATGTCAGAGATCCAA-----CGTGACGAGATTTCATATGAACGG 178
Db 43 CysGluLysHisIleMetGluLysIleGlnGlyArgGlyAspAspAspAspAsp 62
QY 179 GACCCGTACAGCCCTAGTCAGGATCGGTACAGCCCTAGTCCATATGATCGGAGAGCGCT 238
Db 63 AspAspAsnHisIleLeuArgThrMetArgGlyArgIleAsnTyrIleArgArgAsnGlu 82
QY 239 GGATCCTCTCAG-----CACCAGAGAGGTGTTGCAATGAGCTGAAC 280
Db 83 GlyLysAspGluAspGluGluGluGlyHisMetGlnLysCysCysThrGluMetSer 102
QY 281 GAGTTTGAGAAACAACAAAGGTGCATGTCGAGGCAATTCACAGATCATGGAGAACCCAG 340
Db 103 GluLeu---ArgSerProLysCysGlnCysLysAlaLeuGlnLysIleMetGluAsnGln 121
QY 341 AGCGATAGGTTCAGGGGAGGCAACAGAGAGCAACAGTTTCAAGAGGGAGGCTCAGGAACCTT 400
Db 122 SerGluGluLeuGlu---GluLysGlnLysLysLysMetGluLysGluLeuIleAsnLeu 140
QY 401 CCTCAACAGTTCGGCGCTTAGGGCACACAGCGTTTGCAGCTTG 442
Db 141 AlaThrMetCysArgPheGlyProMetIleGlnCysAspLeu 154

RESULT 3
US-10-508-263-42
; Sequence 42, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Glycine max
US-10-508-263-42

Alignment Scores:
Pred. No.: 5.7e-16      Length: 155
Score: 241.50           Matches: 58
Percent Similarity: 57.1%      Conservative: 30
Best Local Similarity: 37.7%   Mismatches: 53
Query Match: 18.9%          Indels: 13
DB: 9                   Gaps: 7

US-10-728-323-2 (1-717) x US-10-508-263-42 (1-155)

QY 2 CTCACCATAGTAGCCCTCGCCCTTTCCTCTCGCTGCCACGCATCTGGAGGAGCAG 61
Db 4 LeuThrIleLeuLeuIleAlaLeuLeuPheIle-----AlaHisThrCysCysAlaSer 21
QY 62 CAGTGGAACTCAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGGCGAACCTGAGG 121
Db 22 LysTrpGlnGlnHisGlnGlnGluSerCysArgGluGlnLeuLysGlyIleAsnLeuAsn 41
QY 122 CCCTGCGAGCAACATCTCATGCAAGATCCAA-----CGTGACGAGATTTCATAT 172
Db 42 ProCysGlu---HisIleMetGluLysIleGlnAlaGlyArgArgGlyGluAspGlySer 60
QY 173 GAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGA 232
Db 61 AspGluAsnHisIleLeuIleArgThrMetPro---GlyArgIleAsnTyrIleArgLys 79
QY 233 GCGCTGGATCCTCTCAG-----CACCAGAGAGGTGTTGCAATGAGCTGAAC 280
Db 80 LysGluGlyLysGluGluGluGluGlyHisMetGlnLysCysCysSerGluMetSer 99

```



```
; Sequence 2, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-508-263-2

Alignment Scores:
Pred. No.: 1,63e-06 Length: 164
Score: 146.50 Matches: 46
Percent Similarity: 46.5% Conservative: 33
Best Local Similarity: 27.1% Mismatches: 52
Query Match: 11.4% Indels: 39
DB: 9 Gaps: 9

US-10-728-323-2 (1-717) x US-10-508-263-2 (1-164)
QY 8 ATACTAGTACCCCTGCCCCCTTTCTCTCGCTGCCCAAGCATCTGCG---AGGCAGCAG 64
   ::::::::::::::::::::
Db 8 ValCysAlaAlaLeuAlaLeuCysPheLeuLeuThrAsnAlaSerIleTyArgThrVal 27
   ::::::::::::::::::::
QY 65 TGGGAACCTCAAGGAGAC-----AGGAAGTCCAG 94
   ::::::::::::::::::::
Db 28 ValGluPheGluGluAspAlaThrAsnProIleGlyProLysMetArgLysCysArg 47
   ::::::::::::::::::::
QY 95 AGCAGCTCGAGAGG---GGAACTCGAGCCCTGCGAGCAACATCTCATGCAGAAAGATC 151
   ::::::::::::::::::::
Db 48 LysGluPheGlnLysGluGlnHisLeuArgAlaCysGlnGlnLeuMetLeuGlnAla 67
   ::::::::::::::::::::
QY 152 CAACGT-----GACGAGGATTATATGAACGGGACCCGTACAGCCCTAGTCAGGAT 202
   ::::::::::::::::::::
Db 68 ArgGlnGlyArgSerAspGluPheAspGluAspMetGluAsnProGln----- 85
   ::::::::::::::::::::
QY 203 CCGTACAGCCCTAGTCATATGATCGGAGAGGCGCTGGATCTCTCAGACCAA----- 256
   ::::::::::::::::::::
Db 86 -----GlyGlnGlnGlnGlnGlnLeu 93
   ::::::::::::::::::::
QY 257 ---GAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAAACCAACAAAGGTGCATGTGCGAG 313
   ::::::::::::::::::::
Db 94 PheGlnGlnCysCysAsnGluLeuArgGlnGluGluProAsp-----CysValCysPro 111
   ::::::::::::::::::::
QY 314 GCATTGCAACAGATCATGGAGAACCGAGAGCATAGGTTGAGGGGAGGCAACAGAGCAA 373
   ::::::::::::::::::::
Db 112 ThrLeuLysGlnAlaAla-----LysAlaValArgLeuGlnGlyGlnHisGlnProMet 129
   ::::::::::::::::::::
QY 374 CAGTTCAAGAGG-----GAGCTCAGAACTTCGCTCAACAGTCGGCCCTTAGGGCA 424
   ::::::::::::::::::::
Db 130 GlnValArgLysIleTyGlnThrAlaLysHisLeuProAsnValCysAspIleProGln 149
   ::::::::::::::::::::
QY 425 CCACAGGCTTGGCAGCTTGGACGTCGAAAGT 454
   ::::::::::::::::::::
Db 150 ValAspValCysProPheAsnIleProSer 159
   ::::::::::::::::::::

RESULT 7
US-11-000-463-715
; Sequence 715, Application US/11000463
; Publication No. US20050266423A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
```

```
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Cao, Yi-Cheng
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785CIP4CN
; CURRENT APPLICATION NUMBER: US/11/000,463
; CURRENT FILING DATE: 2004-11-29
; PRIOR APPLICATION NUMBER: 10/291,265
; PRIOR FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/922,279
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 715
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-000-463-715

Alignment Scores:
Pred. No.: 0.00122 Length: 261
Score: 118.00 Matches: 47
Percent Similarity: 51.1% Conservative: 23
Best Local Similarity: 34.3% Mismatches: 61
Query Match: 9.2% Indels: 6
DB: 11 Gaps: 4

US-10-728-323-2 (1-717) x US-11-000-463-715 (1-261)
QY 460 GCGCCACTTTCGACCTCCAAAGTCGCAACGCTGTGTGCTGCTTAAGCCGCGACATGTTGAGG 401
   ::::::::::::::::::::
Db 12 AlaAlaLeuValLysMetGluLeuLeuValArgTrpAlaProValAlaCysLeuValArg 31
   ::::::::::::::::::::
QY 400 CAAGTTCCTGAGCTCCCTCTTGAACCTGTGCTCTCTGTGCTCCCTCCCTGCAACCTATCGCT 341
   ::::::::::::::::::::
Db 32 GluValAlaLeuGluProLeuAlaLeuLeuValValGluMetMet-----ValLeu 49
   ::::::::::::::::::::
QY 340 CTGGTTCCTCATGATCTGTGTGCAATGCCCTCGCACATGCACCTTTGTTGTTCTC---AAA 284
   ::::::::::::::::::::
Db 50 LeuValLeuProGlyProLeu-----ValProAlaProLeuValLeuAlaSer 67
   ::::::::::::::::::::
QY 283 CTGCTTCAGCTCATTGCAACACACCTCTCTTGGTCTCGAGAGGATCCAGCGCTCTCCGATC 224
   ::::::::::::::::::::
Db 68 LeuValLeuLeuValLeuArgValLysLeuValProLysGlyProGluAlaLeuLysVal 87
   ::::::::::::::::::::
QY 223 ATATGAGCTAGGCTGTACGATCTGTAGCTAGGCTGTAGCGGTGTACGGTCCCGTTTCATATGATC 164
   ::::::::::::::::::::
Db 88 ProArgValCysValValSerLeuAlaProLeuAlaLeuVal---LeuLeuAlaLeu 106
   ::::::::::::::::::::
QY 163 CTCGTCCAGTTGGATCTTCTGCATGAGATGTTCTCGCAGGCTCAGGTTCCGCTCTC 104
   ::::::::::::::::::::
Db 107 LeuGluThrLeuValLeuArgGluSerLeuValLeuLysValProMetValLeuVal 126
   ::::::::::::::::::::
QY 103 GAGCTGCTCTGGCATCTTCTGTCTCTTGGAGTTCCTCAGTTCCTCAGTCTGCTCGC 53
   ::::::::::::::::::::
Db 127 LeuLeuValLeuLeuAlaSerLeuValProGluAlaProLeuAspProArg 143
   ::::::::::::::::::::

RESULT 8
US-11-096-568A-24535
; Sequence 24535, Application US/11096568A
; Publication No. US20060048240A1
```



```
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: Therby
; CURRENT APPLICATION NUMBER: US/11/096,568A
; PRIORITY FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 24535
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(319)
; OTHER INFORMATION: Ceres Seq. ID no. 12435624
US-11-096-568A-24535

Alignment Scores:
Pred. No.: 0.0018 Length: 319
Score: 116.50 Matches: 53
Percent Similarity: 37.7% Conservative: 16
Best Local Similarity: 29.0% Mismatches: 69
Query Match: 9.1% Indels: 45
DB: 11 Gaps: 8

US-10-728-323-2 (1-171) x US-11-096-568A-24535 (1-319)
QY 1 GTCACCATAGTAGAGCCTGCGCTTTTCCTCGCTGCCAGCGATCTGCAGGCA 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 146 AlaProSerAlaSerAlaProThrProCysProSerThrSerProArgAlaCysArgSer 165
QY 61 CGAGTGGGAACCT--CCAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCT 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 166 AlaArgGlySerAlaProArgProProProProProProProProProProSerThrPro 185
QY 118 GAGGCCCTCGAGCAACATCTCATGCGAGAAGATCCAACTGACGAGGATTCATATGAACG 177
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 186 SerSerAlaAlaAla--SerAlaAlaThrThrSerSerSerArgGly-SerSerArgAr 204
QY 178 GGACCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGGCCG 237
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 204 gArgProSerAlaAlaSerThrProThrThrSerProProProSerSerAlaAlaSerSe 224
QY 238 TGGATCCTCTCAGACC-----AAGAGAGGTG 264
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 224 rSerSerAla-AlaThrThrProPheThrSerSerArgPheProSerArgGlnArgSerA 244
QY 265 TTGCATGAGCTGAACGAGT-----TTGAGAACAAACCAAGGTGTCATGTGCGAGGC 315
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 244 laSerThrSerSerTrpSerProThrSerArgProArgProAlaArgCysAlaProA 264
QY 316 ATGC-----AACAGAT 327
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 264 rgCysProGlyArgSerThrSerSerSerThrCysAlaThrProAlaArgGlnArgAr 284
QY 328 CATGGAGACCAAGCAGTAGGTGCGAGGGGACCAAGGACACAGTTCAGAGGGA 387
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 284 erTrpArg-----CysCysArgGlyThrArgAlaSerSerAlaProArg---- 299
QY 388 GCTCAGGAACCTGCTCAACAGTCGCGCTTA-----GGGACCAACAGCGTTGCGA 438
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 300 -----CysArgProThrAlaSerTrpSerProProGlyHisProSerTrpLeua 316
QY 439 CTTGG 443
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 316 laTrp 317

RESULT 9
US-11-033-039-60
; Sequence 60, Application US/11033039
; Publication No. US20060002947A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: LI-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2017US01
; CURRENT APPLICATION NUMBER: US/11/033,039
; PRIORITY FILING DATE: 2005-01-11
; PRIOR APPLICATION NUMBER: 10/245,871
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 1452
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 55
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-11-033-039-55

Alignment Scores:
Pred. No.: 0.00576 Length: 20
Score: 109.00 Matches: 20
```

```
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: LI-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2017US01
; CURRENT APPLICATION NUMBER: US/11/033,039
; PRIORITY FILING DATE: 2005-01-11
; PRIOR APPLICATION NUMBER: 10/245,871
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 1452
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 60
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: hybrid peptide
; FEATURE:
; NAME/KEY: MOD.RES
; LOCATION: (5)..(5)
; OTHER INFORMATION: Ava
US-11-033-039-60

Alignment Scores:
Pred. No.: 0.00479 Length: 25
Score: 110.00 Matches: 21
Percent Similarity: 84.0% Conservative: 0
Best Local Similarity: 84.0% Mismatches: 4
Query Match: 8.6% Indels: 0
DB: 11 Gaps: 0

US-10-728-323-2 (1-171) x US-11-033-039-60 (1-25)
QY 101 CTCGAGAGGGCGAACCTGAGCCCTGCGAGCAACATCTCATGCAAGATCCAACGTGAC 160
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 LeuArgMetLys**LeuArgProCysGluGlnHisLeuMetGlnLysIleGlnArgasp 20
QY 161 GAGGATTCATATGAA 175
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21 GluAspSerTyrGlu 25

RESULT 10
US-11-033-039-55
; Sequence 55, Application US/11033039
; Publication No. US20060002947A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: LI-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2017US01
; CURRENT APPLICATION NUMBER: US/11/033,039
; PRIORITY FILING DATE: 2005-01-11
; PRIOR APPLICATION NUMBER: 10/245,871
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 1452
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 55
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-11-033-039-55

Alignment Scores:
Pred. No.: 0.00576 Length: 20
Score: 109.00 Matches: 20
```

```
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 8.5% Indels: 0
DB: 11 Gaps: 0

US-10-728-323-2 (1-1717) x US-11-033-039-55 (1-20)

QY 116 CTGAGGCCCTCGCAGCAACATCTCATGCGAGAGATCCAAAGTGCAGAGGATTCATATGAA 175
Db 1 LeuArgProCysGluGlnHisLeuMetGlnLysLeuArgAspGluAspSerTyrGlu 20

RESULT 11
US-11-090-908-7
; Sequence 7, Application US/11090908
; Publication No. US20050255574A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Henrik
; APPLICANT: Poulsen, Charlotte Horsmans
; APPLICANT: Soe, Jorn Borch
; APPLICANT: Zargahi, Masoud Rajabi
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: 674509-2011.1
; CURRENT APPLICATION NUMBER: US/11/090,908
; CURRENT FILING DATE: 2005-03-25
; PRIOR APPLICATION NUMBER: PCT/IB2003/004585
; PRIOR FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB 0222512.6
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: GB 0223674.3
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: GB 0228082.4
; PRIOR FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: US 60/431,209
; PRIOR FILING DATE: 2002-12-05
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Triticum sp.
US-11-090-908-7

Alignment Scores:
Pred. No.: 0.0136 Length: 260
Score: 107.50 Matches: 44
Percent Similarity: 40.9% Conservative: 21
Best Local Similarity: 27.7% Mismatches: 60
Query Match: 8.4% Indels: 34
DB: 11 Gaps: 9

US-10-728-323-2 (1-1717) x US-11-090-908-7 (1-260)

QY 1 GCTCACCATACTAGT-----AGCCCTCGCCCTTTCTCTCTCTCTGTC----- 42
Db 102 AlaHisLeuLysSerMetArgMetSerLeuGlnThrLeuProSerMetCysAsnIleTyr 121
QY 43 ---CCAGCATCTCGAGGAGCAGAGTGGGAATCCCAAGGAGACAGAGAAGATCCAGAGCCCA 99
Db 122 ValProValGlnCysGln-GlnGlnGlnGlnLeu-----GlyArgGlnGlnGlnGln 139
QY 100 GCTCGAGAGGGCAACCTGAGGCCCTCGAGCAACATCTCATGCGAAGATCCAAACGTGA 159
Db 139 nLeuGlnGlu---GlnLeuLysProCysAlaThrPheLeuGlnHisGlnCysArgProMe 158
QY 160 CGAGGATTCATATGAACGGACCGGTACAGCCCTAGTCAGGATCCGTCACAGCCCTAGTCC 219
Db 158 tThrValProPhe-----ProHisThrProValGlnLysProThrSer----- 172
QY 220 ATATGATCGGAGGCCCTGATCCTCTCAGCACCAGAGAGGTGTTCGAATGAGCTGAA 279
Db 173 -----CysGlnAsnValGlnSerGlnCysCysArgGlnLeuAl 185
QY 280 CGAGTTTGAACAACCAAGGTGCATGTGCGAGGCATTCGACAGATCATGAGAACCA 339
```

```
Db 185 aGlnIleProGluGlnPheArg-----CysGlnAlaIleHisAsnValValGlu----- 201
QY 340 GAGCGATAGTTTGCAGGGGAGGCAACAGGAGCAACAG-----TTCAA 381
Db 202 -SerIleArgGlnGlnGlnHisGlnProGlnGlnValGlnLeuGluGlyLeuAr 221
QY 382 GAGGAGCTCAGAACTTGCCTCAACAGTGGCGCTTAGGGCACCACACAGCGTTGC 436
Db 221 gMetSerLeuHisThrLeuProSerMetCysLysIleTyrIleProValGlnCys 239

RESULT 12
US-10-508-263-32
; Sequence 32, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SEQ ID NO 32
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Brassica napus
US-10-508-263-32

Alignment Scores:
Pred. No.: 0.0158 Length: 178
Score: 106.50 Matches: 40
Percent Similarity: 41.3% Conservative: 31
Best Local Similarity: 23.3% Mismatches: 56
Query Match: 8.3% Indels: 45
DB: 9 Gaps: 8

US-10-728-323-2 (1-1717) x US-10-508-263-32 (1-178)

QY 8 ATACTAGTAGCCCTCGCCCTTTCTCTCTGTCGCCAGCATCTCG---AGGCAGCAG 64
Db 8 ValSerAlaThrLeuAlaPhePheLeuLeuThrAsnAlaSerIleTyrArgThrVal 27
QY 65 TGGGAACCTCAAGGAGAC-----AGAAGATGC 91
Db 28 valGluPheAspGluAspAlaThrAspSerAlaGlyProPheArgIleProLysCys 47
QY 92 CAGAGCCAGCTCGAGAGGCG---AACCTGAGGCCCTCGCAGCAACATCTCATGCAGAAG 148
Db 48 ArgLysGluPheGlnGlnAlaGlnHisLeuArgAlaCysGlnGlnTrpLeuHisLysGln 67
QY 149 ATCCAACGTGAC-----GAGGAT 166
Db 68 AlaMetGlnSerGlyGlyProSerTrpThrLeuAspGlyGluPheAspPheGluAsp 87
QY 167 TCATATGAACGGACCGGTACAGCCCTAGTCAGAGTCCGTACAGCCCTAGTCCATATGAT 226
Db 88 AspMetGlu---AsnProGlnGlyProGlnGlnArgPro----- 99
QY 227 CGGAGAGGGCGCTGGATCTCTCTCAGACCAAGAGGTGTTCGAATGAGCTGAACGAGTTT 286
Db 100 -----ProLeuLeuGlnGlnCysCysAsnGluLeuHisGlnGlu 112
QY 287 GAGAACCAACAAAGGTGCATGTCGAGGCATTCACACAGATCATGGAGAACCGAGCGCAT 346
Db 113 Gln-----ProLeuCysValCysProThrLeuLysGlyAlaSerLysAlaValLysGln 130
QY 347 AGGTGTCAGGGAGGCAACAGGAGCAACAGTTCAGAGGGAGCTCAGGAACCTTGCCTCAA 406
Db 131 GlnIleGln---GlnGlnGlnGlnGlnGlnGlnGlnGlnMetValSerArgIleTyr 149
QY 407 CAGTGCGGCCTTAGGGCACCAACAGCGTTCGACTTG 442
```

```
Db 150 GlnThrAlaThrHisLeuProLysValCysAsnIle 161
RESULT 13
US-10-508-263-4
; Sequence 4, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-508-263-4
Alignment Scores:
Pred. No.: 0.0195 Length: 164
Score: 105.50 Matches: 40
Percent Similarity: 44.3% Conservative: 30
Best Local Similarity: 25.3% Mismatches: 59
Query Match: 8.2% Indels: 29
DB: 9 Gaps: 9
US-10-728-323-2 (1-717) x US-10-508-263-4 (1-164)
QY 8 ATACTAGTACCCCTCGCCCTTTCTCTCGCTGCGCCAGCATCTGCG---AGGCAGCAG 64
   ::::::::::::::::::::
Db 8 ValCysAlaThrLeuAlaLeuCysPheLeuLeuThrAsnAlaSerIleTyrArgThrVal 27
   ::::::::::::::::::::
QY 65 TGGGAACCTCAGGAGAC-----AGAGATGCCAGAC 97
   ::::::::::::::::::::
Db 28 ValGluPheGluAspAlaSerAsnProValGlyProArgGlnArgCysGlnLys 47
   ::::::::::::::::::::
QY 98 CAGCTCGAGAGGCGC---AACCTGAGCGCCCTGCGAGCAACATCTCATGCAAGATCCAA 154
   ::::::::::::::::::::
Db 48 GluPheGlnSerGlnHisLeuAlaCysGlnArgTrpMetSerLysGlnMetArg 67
   ::::::::::::::::::::
QY 155 COTGACGAGATTATATGAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCT 214
   ::::::::::::::::::::
Db 68 GlnGlyArgGlyGly-----GlyProSerLeuAsp-----77
   ::::::::::::::::::::
QY 215 AGTCCATATGATCGGAGAGCGCTGGATCTCT---CAGCACCAGAGAGGTGTGCAAT 271
   ::::::::::::::::::::
Db 78 AspGluPheAspPheGluGlyProGlnGlnGlyTyrGlnLeuLeuGlnCysCysAsn 97
   ::::::::::::::::::::
QY 272 GAGCTCAACAGATTTCAGAAACCAACAAAGGTGATGTGCGAGGCATTGCAACAGATCATG 331
   ::::::::::::::::::::
Db 98 GluLeuArgGlnGluGlu-----ProValCysValCysProThrLeuLysGlnAlaAla 115
   ::::::::::::::::::::
QY 332 GAGAACACAGAGATAGTTGCGAGGGAGGCAA-----CAGGACCAACAGTTCAAG 382
   ::::::::::::::::::::
Db 116 ArgAlaValSer-----LeuGlnGlyGlnHisGlyProPheGlnSerArgLysIleTyr 133
   ::::::::::::::::::::
QY 383 AGGAGCTCAGGAACTTGCTCAACAGTGGCGCTTAGGGCACCACAGCGTTGC 436
   ::::::::::::::::::::
Db 134 GlnSerAlaLysTyrLeuProAsnIleCysLysIleGlnGlnValGlyGluCys 151
   ::::::::::::::::::::
RESULT 14
US-10-508-263-38
; Sequence 38, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-508-263-38
Alignment Scores:
Pred. No.: 0.0199 Length: 180
Score: 105.50 Matches: 44
Percent Similarity: 39.7% Conservative: 27
Best Local Similarity: 24.6% Mismatches: 49
Query Match: 8.2% Indels: 59
DB: 9 Gaps: 11
US-10-728-323-2 (1-717) x US-10-508-263-38 (1-180)
QY 8 ATACTAGTACCCCTCGCCCTTTCTCTCGCTGCGCCAGCATCTGCG---AGGCAGCAG 64
   ::::::::::::::::::::
Db 8 ValSerAlaThrLeuAlaPhePheLeuLeuThrAsnAlaSerIleTyrArgThrIle 27
   ::::::::::::::::::::
QY 65 TGGGAACCTCAGGAGAC-----AGAGATGC 91
   ::::::::::::::::::::
Db 28 ValGluValAspGluAspAlaThrAsnProAlaGlyProPheArgIleProLysCys 47
   ::::::::::::::::::::
QY 92 CAGAGCCAGCTCGAGAGGCGC---AACCTGAGCGCCCTGCGAGCAACATCTCATGCAAG 148
   ::::::::::::::::::::
Db 48 ArgLysGluPheGlnGlnAlaGlnHisLeuLysAlaCysGlnGlnTrpLeuHisLysGln 67
   ::::::::::::::::::::
QY 149 ATCCAACTGAC-----GAGGAT 166
   ::::::::::::::::::::
Db 68 AlaMetGlnSerGlySerGlyProSerTrpThrLeuAspGlyGluPheAspPheGluAsp 87
   ::::::::::::::::::::
QY 167 TCATATGAGCGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGAT 226
   ::::::::::::::::::::
Db 88 AspMetGlu---AsnProGlnGlyProGlnGlnArgPro-----99
   ::::::::::::::::::::
QY 227 CGGAGAGCGCTGGATCTCTCAGCACCACCAAGAGAGGTGTGCAATGAGTGAACGAGTTT 286
   ::::::::::::::::::::
Db 100 -----ProLeuLeuGlnGlnCysCysAsnGlnLeuHisGlnGlu 112
   ::::::::::::::::::::
QY 287 GAGACAACCAAGTGTGATGTCGAGGCAATTG-----CAA 322
   ::::::::::::::::::::
Db 113 Glu-----ProLeuCysValCysProThrLeuLysGlyAlaSerLysAlaValLysGln 130
   ::::::::::::::::::::
QY 323 CAGATCATGGAACACAGAGCATAGTTGCGAGGGAGGACAGCAGGACAGTTCAG 382
   ::::::::::::::::::::
Db 131 GlnValArgGlnGlnGlnGln-----GlnGlyGlnGlnLeuGlnGlnValIleSer 148
   ::::::::::::::::::::
QY 383 AGG-----GAGCTCAGGAACTTGCTCAACAGTGGCGCTTAGGGCACCACAC 430
   ::::::::::::::::::::
Db 149 ArgIleTyrGlnThrAlaThrHisLeuProLysValCysAsnIle-----ProGln 165
   ::::::::::::::::::::
RESULT 15
US-10-508-263-40
; Sequence 40, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Sinapis alba
US-10-508-263-40
Alignment Scores:
Pred. No.: 0.0269 Length: 144
```

```
Score: 104.00 Matches: 37
Percent Similarity: 42.0% Conservative: 23
Best Local Similarity: 25.9% Mismatches: 47
Query Match: 8.1% Indels: 36
DB: 9 Gaps: 7

US-10-728-323-2 (1-717) x US-10-508-263-40 (1-144)

QY 77 GGAGAGAGAGATGCCAGAGCCAGCTCGAGAGCGCG---AACCTGAGGCCCTCGAGAGCAA 133
Db 6 GlyIleProLysCysArgLysGluPheGlnGlnAlaGlnHisLeuArgAlaCysGlnGln 25
QY 134 CATCTCATGCGAGAGATCCAACTG-----GACGAG 163
Db 26 TrpLeuHisLysGlnAlaMetGlnSerGlySerGlyProSerTrpThrLeuAspAspGlu 45
QY 164 GATTCTATATGAACGGGACCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATAT 223
Db 46 PheAspPheGluAspMetGluAsnPro---GlnGlyProGlnGlnArgProProLeu 64
QY 224 GATCGGAGAGCGCTGGATCCTCTCAGCACCAAGAGAGAGGTGTTGCAATGAGCTGAACGAG 283
Db 65 -----LeuGlnGlnCysCysAsnGluLeuHisGln 74
QY 284 TTTGAGNACAAACAGGTGCTGCGAGGCATTGCAACAGATCATGGAG-----AAC 337
Db 75 GluGlu-----ProLeuCysValCysProThrLeuLysGlyAlaSerLysAlaValLys 92
QY 338 CAGAGCGATAGGTTGCGAGGGGAGGCAACAGGAGCAACAG----- 376
Db 93 GlnGlnValArgGlnGlnLeuGluGlnGlnGlnGlnGlnGlnGlnGlnGlnHisVal 112
QY 377 -----TTCAGAGGGAGCTCAGGAATTGCTCAACAGTCCGGCCTTAGGGCACCA 427
Db 113 IleSerArgIleTyrGlnThrAlaThrHisLeuProArgValCysAsnIleArgGlnVal 132
QY 428 CAGCGTTGC 436
Db 133 SerValCys 135
```

Search completed: May 16, 2006, 00:03:03
Job time : 15.3331 secs

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: May 15, 2006, 21:46:12 ; Search time 4.79902 Seconds
(without alignments)
4312.597 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 1281
Sequence: 1 gctcaccatactagtagcccc.....taaaagaatcatgttttgg 717

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abs/ABSWEB/spool/US10728323/runat_15052006_172134_22413/app_query.fasta_1
-DB=PIR -QFMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -MODE=LOCAL
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p
-USER=US10728323 @CN1.1.92 @runat_15052006_172134_22413 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	278	21.7	153	A33090	conglutin delta pr
2	243	19.0	158	T05710	2S albumin precurs
3	187.5	14.6	80	B23617	conglutin delta-2
4	184.5	14.4	323	S38887	2S albumin - commo
5	173.5	13.5	141	T10257	2S albumin precurs
6	154	12.0	154	S14947	2S albumin - Brazil
7	151	11.8	295	S01062	2S seed storage pr
8	147	11.5	186	A29802	napin precursor (g
9	146.5	11.4	164	NW01	2S albumin 1 precu
10	140.5	11.0	146	S14946	2S seed storage pr
11	139	10.9	155	JCS379	mablin II precur
12	139	10.9	258	RZCS	2S seed storage pr
13	132.5	10.3	165	T09252	seed storage prote
14	129	10.1	170	T08012	2S seed storage pr

15	128	10.0	172	2	S18871	2S-like storage pr
16	126.5	9.9	70	2	A59346	seed storage prote
17	125	9.8	162	2	T08013	2S seed storage pr
18	124.5	9.7	162	2	S49259	albumin 4 - easter
19	122.5	9.6	162	2	T08010	2S seed storage pr
20	121	9.4	100	2	S48180	mablin IV - Yun
21	121	9.4	174	2	P50425	napin B33 precursor
22	119.5	9.3	180	2	S20205	napin (clones BNMN
23	118.5	9.3	106	2	S26636	napin n1b - rape
24	118.5	9.3	110	2	S20350	napin n1a - rape
25	118	9.2	104	2	S48176	mablin I-1 - Yun
26	117	9.1	104	2	S48178	mablin III - Yun
27	115	9.0	152	2	P50427	napin AHI precursor
28	111.5	8.7	173	2	T08011	2S seed storage pr
29	108.5	8.5	178	2	S07828	napin B - rape
30	108.5	8.5	178	2	S25127	2S storage protein
31	107	8.4	139	2	T09850	albumin 2S storage
32	107	8.4	139	2	T09878	albumin 2S storage
33	106.5	8.3	178	1	NWRP2	napin 2 precursor
34	106.5	8.3	178	2	A25997	napin precursor (n
35	105.5	8.2	164	1	NW03	2S albumin 3 precu
36	105.5	8.2	180	2	S10018	napin (clone BNGA
37	105.5	8.2	1099	2	A56155	tumor suppressor p
38	105.5	8.2	1909	2	A45592	liver stage antige
39	104.5	8.2	1263	2	T15496	hypothetical prote
40	104	8.1	145	2	S65479	allergen Sin a I (
41	103.5	8.1	166	1	NW04	2S albumin 4 precu
42	103.5	8.1	178	2	S25130	2S storage protein
43	102.5	8.0	178	2	S25134	2S storage protein
44	101.5	7.9	178	2	S25137	2S storage protein
45	101	7.9	145	2	PC1246	Sin a I allergen 1

ALIGNMENTS

RESULT 1

A33090
conglutin delta precursor - narrow-leaved blue lupine
C:Species: Lupinus angustifolius (narrow-leaved blue lupine)
C>Date: 03-Aug-1990 #sequence_revision 03-Aug-1990 #text_change 09-Jul-2004
C:Accession: S12404; A33090
R:Gayler, K.R.; Kolivas, S.; Macfarlane, A.J.; Lilley, G.G.; Baldi, M.; Blagrove, R.J.; Plant Mol. Biol. 15, 879-893, 1990
A:Title: Biosynthesis, cDNA and amino acid sequences of a precursor of conglutin delta,
A:Reference number: S12404; MUID:91355912; PMID:2103479
A:Accession: S12404
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-153 <GAY>
A:Cross-references: UNIPROT:Q99235; UNIPARC:UPI000009D797; GB:X53523; NID:g19140; PIDN:
C:Superfamily: soybean 2S albumin
C:Keywords: seed
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-59/Product: conglutin small chain #status predicted <MAT1>
F:73-153/Product: conglutin large chain #status predicted <MAT2>

Alignment Scores:	Pred. No.:	Score:	Length:	Matches:
			3 53e-18	153
			278.00	62
Percent Similarity:	55.0%	Conservative:	26	
Best Local Similarity:	38.8%	Mismatches:	46	
Query Match:	21.7%	Indels:	26	
DB:	2	Gaps:	5	
US-10-728-323-2 (1-1717) x A33090 (1-153)				

QY	2	CTCACCATACTAGTAGCCCTC---GCCCTTTTCTCTCGTCCGCCACCATCTGCGAGG 58
Db	4	LeuThrIleuIleAlaLeuValAlaAlaLeuValLeuValHisThrSerAlaPhe 23
QY	59	GAGCAGTGGGAACCTCCAGGACACAGACATCCAGACCTCGAGAGGGCGACCTTG 118
Db	24	Gln-----SerSerLysGlnSerCysLysArgGlnLeuGlnValAsnLeu 39

```
QY 119 AGCCCTCGCAGCAACATCTCATGACAGATCCAAACGTGACGAGGATTCATATGAACGG 178
Db 40 ArgHisCysGluAsnHisIleAlaGlnArgIleGlnGlnGluGluGluAsp 59
QY 179 GACCCGTACAGC-----CCTAGTCAGGATCCG 205
Db 60 HisAlaLeuLysLeuArgGlyIleLysHisValIleLeuArgHisArgSerGlnGlu 79
QY 206 TACAGCCCTAGTCCATATGATCGAGAGCGCTGGATCCTCTCAGCACCAAGAGAGGTGT 265
Db 80 TyrSerGlu-----GluSerGluGluLeuAspGlnCys 90
QY 266 TGAATGAGCTGAACGAGTTTGAGAACCAAAAGGTGATGTCGAGGCAATTCGCAACAG 325
Db 91 CysGluGluLeuAsnGluLeu---AsnSerGlnArgCysGlnCysArgAlaLeuGlnGln 109
QY 326 ATCATGGAACAGAGCATAGTTGCGAGGGAGGCAACAGGAGCAACAGTTCAGAGG 385
Db 110 IleTyrGluSerGlnSerGlnGlnCysGluGlySerGlnGlnGlnGlnGluGlnGln 129
QY 386 GAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGGGCCACACAGCGTTGCGCACTTGGAC 445
Db 130 GluLeuGluLysLeuProArgThrCysGlyPheGlyProLeuArgArgCysAspValAsn 149

RESULT 2
T05710
2S albumin precursor - soybean
N;Alternate names: aspartic acid-rich peptide
C;Species: Glycine max (soybean)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: T05710; A28485
R;Wang, J.; Pichersky, E.
Plant Physiol. 114, 1567, 1997
A;Title: Nucleotide sequence of S-Adenosyl-L-Methionine:Caffeic Acid 3-O-Methyltransferase
A;Reference number: Z15424
A;Accession: T05710
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-158 <WNA>
A;Cross-references: UNIPROT:P19594; UNIPARC:UPI00000005F1; EMBL:AF005030; NID:g2305019;
A;Experimental source: cultivar Hodgson 78; cotyledon
R;Odani, S.; Koide, T.; Ono, T.
J. Biol. Chem. 262, 10502-10505, 1987
A;Title: Amino acid sequence of a soybean (Glycine max) seed polypeptide having a poly(U
A;Reference number: A28485; MUID:87280104; PMID:3611081
A;Accession: A28485
A;Molecule type: protein
A;Residues: 22-64 <ODA>
A;Cross-references: UNIPARC:UPI0000003C545
C;Superfamily: soybean 2S albumin
C;Keywords: glycoprotein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-64/Product: 2S albumin small chain #status predicted <NAT>
F;54-56/Region: cell attachment (R-G-D) motif
F;39/Binding site: carbonylate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 7,31e-15 Length: 158
Score: 243.00 Matches: 56
Percent Similarity: 56.5% Conservative: 31
Best Local Similarity: 36.4% Mismatches: 55
Query Match: 19.0% Indels: 12
DB: 1 Gaps: 6

US-10-728-323-2 (1-717) x T05710 (1-158)

QY 5 ACCATAGTAGCCCTCGCCCTTTCTCCTCGCTGCCACGCAATCGAGGAGCAG 64
Db 5 ThrIleLeuLeuIleSerLeuPheCysIle---AlaHisThrCysSerAlaSerLys 23
QY 65 TGGGAACCTCCAGGAGACAGAAATGCGAGCGCAGCTCGAGAGGGCGAACCTTGAGGCC 124
Db 65 TGGGAACCTCCAGGAGACAGAAATGCGAGCGCAGCTCGAGAGGGCGAACCTTGAGGCC 124

Alignment Scores:
Pred. No.: 7.31e-15 Length: 158
Score: 243.00 Matches: 56
Percent Similarity: 56.5% Conservative: 31
Best Local Similarity: 36.4% Mismatches: 55
Query Match: 19.0% Indels: 12
DB: 1 Gaps: 6

US-10-728-323-2 (1-717) x B23617 (1-80)

QY 245 TCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAACACCAAGTGTC 304
Db 11 SerGluGluLeuAspGlnCysCysGluGlnGlnLeuAsnGlnLeu---AsnSerGlnArgCys 29
QY 305 ATGTGCGAGGCATTGCAACAGATCATGGAGAACACGAGCGATAGGTTCAGGGAGGCA 364
Db 30 GlnCysArgAlaLeuGlnGlnIleTyrGluSerGlnSerGlnCysGluGlyArgGln 49
QY 365 CAGGAGCAACAGTTCAAGAGGAGGCTCAGGAACCTTCCTCAACAGTTCGCGCTTAGGGCA 424
Db 50 GlnGluGlnGlnLeuGluGlyGluLeuGluLysLeuProHargileCysGlyPheGlyPro 69
QY 425 CCACAGCGTTGCGACTTGGAC 445
Db 70 LeuArgArgCysAsnIleAsn 76

RESULT 4
S38887
2S albumin - common sunflower (fragment)
C;Species: Helianthus annuus (common sunflower)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
```

```
Db 24 TrpGlnHisGlnGlnAsp---SerCysArgLysGlnLeuGlnGlyValAsnLeuThrPro 42
QY 125 TCGAGCAACATCTCATGCAAGATCCAA-----CGTGACGAGGATTCATATGAACGG 178
Db 43 CysGluLysHisIleMetGluLysIleGlnGlyArgGlyAspAspAspAspAsp 62
QY 179 GACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGGCGCT 238
Db 63 AspAspAsnHisIleLeuArgThrMetArgGlyArgIleAsnTyrIleArgArgAsnGlu 82
QY 239 GGATCTCTCTCAG-----CACCAAGAGAGGTGTTGCAATCAGCTGAAC 280
Db 83 GlyLysAspGluAspGluGluGluGlyHisMetGlnLysCysCysThrGluMetSer 102
QY 281 GAGTTTGAGAACCAACAAAGGTGCTATGTCGAGGCAATTCGAACAGATCATCGAGAACCCAG 340
Db 103 GluLeu---ArgSerProLysCysGlnCysLysAlaLeuGlnLysIleMetGluAsnGln 121
QY 341 ACGGATAGTTGCGAGGGAGGCAACAGGAGCAACAGTTCGAAGAGGAGCTCAGGAAGTTG 400
Db 122 SerGluGluLeuGlu---GluLysGlnLysLysLysMetGluLysGluLeuIleAsnLeu 140
QY 401 CCTCAACAGTGGCGGCTTAGGGCCACACAGCGTTGCGACTTG 442
Db 141 AlaThrMetCysArgPheGlyProMetIleGlnCysAspLeu 154

RESULT 3
B23617
conglutin delta-2 large chain - narrow-leaved blue lupine
C;Species: Lupinus angustifolius (narrow-leaved blue lupine)
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 09-Jul-2004
C;Accession: B23617
R;Lilley, G.G.; Inglis, A.S.
FEBS Lett. 195, 235-241, 1986
A;Title: Amino acid sequence of conglutin delta, a sulfur-rich seed protein of Lupinus
A;Reference number: A91358
A;Accession: B23617
A;Molecule type: protein
A;Residues: 1-80 <LIL>
A;Cross-references: UNIPROT:P09931; UNIPARC:UPI00000127596
C;Superfamily: soybean 2S albumin

Alignment Scores:
Pred. No.: 1.35e-09 Length: 80
Score: 187.50 Matches: 34
Percent Similarity: 73.1% Conservative: 15
Best Local Similarity: 50.7% Mismatches: 17
Query Match: 14.6% Indels: 1
DB: 2 Gaps: 1

US-10-728-323-2 (1-717) x B23617 (1-80)

QY 245 TCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAACACCAAGTGTC 304
Db 11 SerGluGluLeuAspGlnCysCysGluGlnGlnLeuAsnGlnLeu---AsnSerGlnArgCys 29
QY 305 ATGTGCGAGGCATTGCAACAGATCATGGAGAACACGAGCGATAGGTTCAGGGAGGCA 364
Db 30 GlnCysArgAlaLeuGlnGlnIleTyrGluSerGlnSerGlnCysGluGlyArgGln 49
QY 365 CAGGAGCAACAGTTCAAGAGGAGGCTCAGGAACCTTCCTCAACAGTTCGCGCTTAGGGCA 424
Db 50 GlnGluGlnGlnLeuGluGlyGluLeuGluLysLeuProHargileCysGlyPheGlyPro 69
QY 425 CCACAGCGTTGCGACTTGGAC 445
Db 70 LeuArgArgCysAsnIleAsn 76

RESULT 4
S38887
2S albumin - common sunflower (fragment)
C;Species: Helianthus annuus (common sunflower)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
```


Alignment Scores:		1.99e-06	Length: 154
Pred. No.:	Score:	154.00	Matches: 45
Percent Similarity:		45.3%	Conservative: 28
Best Local Similarity:		28.0%	Mismatches: 52
Query Match:		12.0%	Indels: 36
DB:		2	Gaps: 10
US-10-728-323-2 (1-717) x S14947 (1-154)			
QY	8	ATACTAGTACCTCGCCCTTTCTCTCGTCCCGCAGCATCTGCG	----- 55
DB	7	ValAlaAlaLeuLeuAlaLeuLeuValLeuGlyGlnAlaThrAlaPheArgThrThr	26
QY	56	-----AGGCAGCAGTGGGAACCTCAAGGA	-----GACAGAAGA 88
DB	27	ValThrThrThrLeuGluGluGluGluGluAsnProArgGlyArgSerGluGlnGln	46
QY	89	TGCAGAGCCAGCTCGAGG-----GCGAAGCTGAGCCCTGCGAGCAACATCTCATGCGAG	145
DB	47	CysArgGluGlnMetGluArgGlnGlnGlnLeuAsnHisCysArgMetTyrLeuArgGln	66
QY	146	AAGATCCAAAGTGCAGGAGTTCATATGAACGGGACCGGTACAGCCCTAGTCAGGATCCG	205
DB	67	GlnMet-----GluGluSerProTyrGln-----	74
QY	206	TACAGCCCTAGTCCATATGATCGGAGGCGCTGGATCTCTCAGCACCAAGAGAGGTGT	265
DB	75	---AsnProArgProLeuArgArg-----GlyGluGluProHisLeuAspGluCys	90
QY	266	TGCAATGAGCTGAACGAGTTTGAGAACCAACCAAGAGTGCATGTGCGAGGCATTCGACACAG	325
DB	91	CysGluGlnLeuGluArgMetAsp-----GluMetCysArgCysGluGlyLeuArgMet	108
QY	326	ATCATGGAGAACAGAGCGAT-----AGCTTGCGGGGAGGCAACAGGAGCAACAGTTC	379
DB	109	MetLeuArgArgGlnArgGluGluMetGluLeuGlnGly---GluGlnMetGlnArgIle	127
QY	380	AAGAGGGAGCTCAGGAATTCCTCAACAGTGGCGCTTAGGGCACACACAGCTTTCGAC	439
DB	128	MetArgGlyAlaGluAsnLeuLeuSerArgCysAsnLeu---SerProGlnArgCysPro	146
QY	440	TTG 442	
DB	147	Met 147	
RESULT 7			
S01062			
2S seed storage protein precursor (clone HaGS) - common sunflower			
N;Alternate names: 2S albumin storage protein			
C;Species: Helianthus annuus (common sunflower)			
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004			
C;Accession: S01062			
R;Allen, R.D.; Cohen, E.A.; Vonder Haar, R.A.; Adams, C.A.; Ma, D.P.; Nessler, C.L.; Th			
Mol. Gen. Genet. 210, 211-218, 1987			
A;Title: Sequence and expression of a gene encoding an albumin storage protein in sunfl			
A;Reference number: S01062; MUID:88142538; PMID:2830455			
A;Accession: S01062			
A;Molecule type: DNA			
A;Residues: 1-295 <ALL>			
A;Cross-references: UNIPROT:P15461; UNIPARC:UPI0000124EBB; EMBL:X06410; NID:g18799; PIDN			
A;Note: part of this sequence, including the amino end of the mature protein, was confi			
C;Genetics:			
F;1-20/Domain: signal sequence #status predicted <SIG>			
F;162-295/Product: 2S seed storage protein #status experimental <MAT>			
Alignment Scores:			
Pred. No.:	Score:	3.77e-06	Length: 295
Percent Similarity:		151.00	Matches: 47
Best Local Similarity:		40.4%	Conservative: 35
		23.2%	Mismatches: 65

Query Match:		11.8%	Indels: 56
DB:		2	Gaps: 8
US-10-728-323-2 (1-717) x S01062 (1-295)			
QY	8	ATACTAGTACCTCGCCCTTTCTCTCGTCCCGCAGCATCTGCG	----- 55
DB	5	IleValLeuAlaLeuAlaPheAlaLeuValAlaPheAlaThrAlaHisThrIle	24
QY	56	-----AGGCAGCAGTGGGAACCTCAAGAGAGACAGACAGATGCCAGACCCAG	100
DB	25	IleThrThrThrIleGluAspGluAsnProIleSerGlyGlnArgGlnValSerGlnArg	44
QY	101	CTCGAGAGGCGCAACCTGAGGCCCTCGGAGCAACATCTCTCATG-----CAGAAGATCCAA	154
DB	45	IleGlnGlyGlnArgLeuAsnGlnCysArgMetPheLeuGlnGlnGlyGlnAsnIlePro	64
QY	155	CGTGCAGGAGTTCATATGAACGGGACCGGTACAGCCCTAGTCAGGATCCGTACAGCCCT	214
DB	65	ArgGluPheAspAsn-----	69
QY	215	AGTCCATATGATCGGAGAGGCGCTGGATCCTCTCAGCACCAAGAGAGGTGTTCGAATGAG	274
DB	70	-----ProGlnMetGlyArgGlnGlnGlnGlnLeuGlnGlnCysCysGlnGlu	86
QY	275	CTGACAGCTTTGAGAACCAAGGTGCATGTGCGAGGCATTCGACAGATCATCGGAG	334
DB	87	LeuGlnAsnIleGly-----GlnCysGlnCysGluAlaValIleValPheArg	104
QY	335	AACACAGGCGATAGTTGCAG-----GGGAGG	361
DB	105	GluAlaGlnGlnValGlnGlnGlnGlnGlnGlnGlnValProPheArgGlySer	124
QY	362	CRACAGAGCAACAGTTCAAGAGGAGGAGTTCCTCAACAGTTCGCGCCCTTAGG	421
DB	125	GlnGlnThrGlnGlnLeuLysGlnLysAlaGlnIleLeuProAsnValCysAsnLeuGln	144
QY	422	GCACACAGCTTGCACCTTG-----GACGTCGAA	451
DB	145	Ser---ArgArgCysGluIleGlyThrThrThrValThrGluSerAsnIleAsp	163
QY	452	AGTGGCGGAGACAGACAGATCTAAACACCTATCTCAAAAAAGAGAAAGAAAGAA	511
DB	164	IleProPheArgAspArg-----ProPheGlyThrGlySerGlnGlnCysArgGlu	180
QY	512	AATAGCTTA 520	
DB	181	ThrGluIle 183	
RESULT 8			
A29802			
N;Alternate names: 1.7 S seed storage protein			
C;Species: Brassica napus (rape)			
C;Date: 20-Jun-1989 #sequence_revision 20-Jun-1989 #text_change 31-Dec-2004			
C;Accession: A29802			
R;Scofield, S.R.; Crouch, M.L.			
J. Biol. Chem. 262, 12202-12208, 1987			
A;Title: Nucleotide sequence of a member of the napin storage protein family from Brass			
A;Reference number: A29802; MUID:87308225; PMID:3040733			
A;Accession: A29802			
A;Molecule type: DNA			
A;Residues: 1-186 <SCO>			
A;Cross-references: UNIPROT:P09893; UNIPARC:UPI0000124EBE; GB:J02782; NID:g167178; PIDN			
C;Genetics:			
A;Gene: gNA			
C;Superfamily: Alpha amylase inhibitor			
Alignment Scores:			
Pred. No.:	Score:	9.11e-06	Length: 186
Percent Similarity:		147.00	Matches: 47
Best Local Similarity:		41.0%	Conservative: 28
		25.7%	Mismatches: 52

Db 63 -----GluSerProTyr-GlnThrMet 69

Qy 218 CCATATGATCGGAGAGCGGTGGATCCTCTCAGCACCAAGAGAGGTGTTCGAATGACGTG 277
||| ||||| ||| ||| ||| ||| ||| |||
Db 70 Pro-----ArgArg-----GlyMetGluProHisMetSerGluCyS-CysGluGlnLeu 85
||| ||| ||| ||| ||| ||| ||| |||

Qy 278 AACGAGTTTCAGAAACAACCAAGAGTGTCATGTGCAGGCAATTGCCAA---CAGATCATGGAG 334
||| ||| ||| ||| ||| ||| ||| |||
Db 86 GluGlyMetAsp-----GluSerCysArgCysGluGlyLeuA-rgMetMetMetMetArg 103
||| ||| ||| ||| ||| ||| ||| |||

Qy 335 AACCAGAGCGCATAGTTTCAGGGGAGG---CAACAGGAGCAACAGTTCACAGAGGAGTC 391
||| ||| ||| ||| ||| ||| ||| |||
Db 104 MetGlnGlnGluMetGlnProA-rgGlyGluGlnMetA-rgMetMetA-rgLeuAla 123
||| ||| ||| ||| ||| ||| ||| |||

Qy 392 AGGNACTTGCTCAACAGTAGCGGCTTAGGGACCACACAGCGCTTCGCACTTG 442
||| ||| ||| ||| ||| ||| ||| |||
Db 124 GluAsnIleProSerA-rgCysAsnLeu---SerProMetA-rgCysProMet 139
||| ||| ||| ||| ||| ||| ||| |||

RESULT 11

mabinlin II precursor - Yunnan caper

C;Species: Caparis masaikai (Yunnan caper)

C;Date: 02-Jun-1997 #sequence_revision 12-Sep-1997 #text_change 31-Dec-2004

C;Accession: J05379; PC4316

R;Nirasawa, S.; Masuda, Y.; Nakaya, K.; Kurihara, Y.

Gene 181, 225-227, 1996

A;Title: Cloning and sequencing of a cDNA encoding a heat-stable sweet protein.

A;Reference number: J05379; MUID:97128796; PMID:8973336

A;Accession: J05379

A;Molecule type: mRNA

A;Residues: 1-155 <NR1>

A;Cross-references: UNIPROT:P30233; UNIPARC:UPI000016DD17; DDBJ:D83997; NID:g1

A;Accession: PC4316

A;Molecule type: protein

A;Residues: 36-41,149-154 <NR2>

A;Cross-references: UNIPARC:UPI00001763F3; UNIPARC:UPI00001763F4

A;Experimental source: seed

C;Keywords: sweet taste

F;1-20/Domain: signal sequence #status predicted <SIG>

F;21-35/Domain: amino-terminal propeptide #status predicted <PRO>

F;36-68,83-154/Product: mabinlin II #status experimental <MAT>

F;36-68/Domain: mabinlin II A chain #status experimental <CHA>

F;69-83/Domain: mabinlin linker peptide #status predicted <LNK>

F;83-154/Domain: mabinlin II B chain #status experimental <CHB>

Alignment Scores:		
Pred. No.:	5,24e-05	Length: 155
Score:	139.00	Matches: 45
Percent Similarity:	45.5%	Conservative: 26
Best Local Similarity:	28.8%	Mismatches: 47
Query Match:	10.9%	Indels: 38
DB:	2	Gaps: 9

US-10-728-323-2 (1-717) x J05379 (1-155)

Qy 8 ATACTAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCACTCTCGCAGG----- 58
||| ||||| ||||| ||| ||| ||| ||| |||

Db 7 LeuPheAlaThrLeuAlaLeuPheValLeuLeuAlaAsnAlaSerIleGlnThrVal 26
||| ||| ||| ||| ||| ||| ||| |||

Qy 59 -----CAGCAGTGGGAACCTCCAAGGAGACAGAAGATGC 91
||| ||| ||| ||| ||| ||| ||| |||

Db 27 IleGluValAspGluGluAspAsnGlnLeuTrp-----A-rgCys 40
||| ||| ||| ||| ||| ||| ||| |||

Qy 92 CAGAGCCAG-----CTCGAGAGGGCGNAACCTGTAGGCCCTCGGAGCAACATCTCATGCAGAG 148
||| ||| ||| ||| ||| ||| ||| |||

Db 41 GlnA-rgGlnPheLeuGlnHisGlnA-rgLeuLeuAlaCysGlnA-rgPheIleHisA-rgA-rg 60
||| ||| ||| ||| ||| ||| ||| |||

Qy 149 ATCCAA-----CGTGACGAGGATTCATATGAACGGGACCCGTACAGCCCTAGTCAGGAT 202
||| ||| ||| ||| ||| ||| ||| |||

Db 61 AlaGlnPheGlyGlyGlnProA-rgGluLeuGluA-rgValGluA-rgAspAsnA-rgA-rg 80
||| ||| ||| ||| ||| ||| ||| |||

Qy 203 CCGTACAGCCCTAGTCCATATGATTCGGAGAGCGCGTGGATTCCTCTCAGCACCAAGAGAGG 262
||| ||| ||| ||| ||| ||| ||| |||


```
Query Match: 10.3% Indels: 65
DB: 2 Gaps: 10

US-10-728-323-2 (1-717) x T09252 (1-165)
QY 8 ATACTAGTAGCCTCGCCCTTTCTCTCGCT-----GCC 43
   ::: ::: :|||||
Db 17 ValSerLeuGlyMetAlaLeuLeuLeuLeuLeuHisTrpGlyThrArgThrValAspAla 36
   |||
QY 44 CAGCCTCTCCGAGGCGAGCTGGAACTCCAGGAGACAGAGA---TGCCAGAGCCAG 100
   |||
Db 37 HisGluAspGlyLeuTyrGlyGluValGlnGlnArgSerCysGluGlnGln 56
   |||
QY 101 CTCGAGGCGGCAACCTGAGGCCCTCGGACCAACATCTCATGCAAGATCCCAACGTGAC 160
   |||
Db 57 -----ArgLeuSerSerCysArgGluTyrLeu-----GluArgPro 68
   |||
QY 161 GAGGATTATGAAACGGGACCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCA 220
   |||
Db 69 ArgAsp----- 70
QY 221 TATGATCGGAGAGCGCTGGATCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAAC 280
   |||
Db 71 -----GlnProSerGluArgCysGluGluLeuGln 81
QY 281 GAGTTTGAGAACCAACAGGTGCGAGGCATTGCAACAGATCATGGAAGAAC--- 337
   |||
Db 82 ArgMet-----SerProGlnCysArgCysGlnAlaIleGlnArgThrLeuGluAspVal 99
   |||
QY 338 -----CAGAGCGAT-----AGG 349
   |||
Db 100 PheMetAspSerAspSerGlnAspGlyAlaProLeuAsnGlnArgArgArgGlnArgArg 119
QY 350 TTCAGGGGAGG---CAACAGGAGCAACAGTTCAAGAGGAGGCTCAGGAACCTTCCTCAA 406
   |||
Db 120 GlyGlnGlyArgGlyMetGluGluGluValValArgAlaGluGluLeuProAsn 139
QY 407 CAGTGGCGCTTAGG---GCACCACAGCGTTGGACTTGCAGCTCGGAAAGT 454
   |||
Db 140 ThrCysAsnValArgGlnSerProArgArgCysAspLeuGlnArgHisSer 156

RESULT 14
T08012
2S seed storage protein precursor - Douglas fir
C;Species: Pseudotsuga menziesii (Douglas fir)
C;Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 31-Dec-2004
C;Accession: T08012
R;Chatthai, M.; Mista, S.
Planta 206, 138-145, 1998
A;Title: Structure and expression of embryogenesis-specific cDNAs encoding 2S seed storage
A;Reference number: Z16286; MUID:98381325; PMID:9715536
A;Accession: T08012
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-170 <CHA>
A;Cross-references: UNIPROT:O64931; UNIPARC:UPI000000AA754; EMBL:AF029972; NID:G2988481;
C;Genetics:
A;Gene: 2S3
C;Superfamily: 2S albumin-like seed storage protein
C;Keywords: seed; storage protein

Alignment Scores:
Pred. No.: 0.000464 Length: 170
Score: 129.00 Matches: 43
Percent Similarity: 36.1% Conservative: 22
Best Local Similarity: 23.9% Mismatches: 49
Query Match: 10.1% Indels: 66
DB: 2 Gaps: 8

US-10-728-323-2 (1-717) x T08012 (1-170)
QY 8 ATACTAGTAGCCTCGCCCTTTCTCTCGCT-----GCCAGCAGTGG 67
   ::: ::: :|||||

US-10-728-323-2 (1-717) x S18871 (1-172)
QY 14 GTAGCCCTCGCCCTTTCTCTCTCTC-----GCTGCCACGCA 49
   ::: ::: :|||||
Db 18 LeuSerValAlaLeuPheLeuLeuPheHisTrpGlyLeProSerValAspGlyHisGlu 37
QY 50 TCTCGAGGCGAGCTGGGAATCCAAAGGAGACAGAGA---TGCCAGAGCCAGCTCGAG 106
   |||
Db 38 AspAsnMetTyrGlyGluGluIleGlnGlnArgSerCysAspProGlnArgAsp 57
   |||
QY 107 AGGCGCAACCTGAGGCCCTCGGACCAACATCTCATGCAAGATCCCAACGTGACGAGAT 166
   |||
Db 58 ProGlnArgLeuSerSerCysArgAspTyrLeu----- 68
QY 167 TCATATGAACGGGACCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGAT 226
```

```

Db      69  -----GluArgArgGluGlnProSer----- 76
QY      227  CGGAGAGCGCTGGATCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAGCGAGTTT 286
Db      77  -----GluArgCysCysGluGluLeuGlnArgMet 86
QY      287  GAGAACAAACAAAGGTGCATGTGCGAGGCATTGCAACAGATCATGGAG----- 334
Db      87  -----SerProGlnCysArgCysGlnAlaIleGlnMetLeuAspGlnSerLeuSer 104
QY      335  -----AACCAGAGCGATAGG 349
Db      105  TyrAspSerPheMetAspSerAspSerGlnGluAspAlaProLeuAsnGlnArgArg 124
QY      350  TTG-----CAGGGAGGCAACAGAGCAACAGTTTCAAGAGGAGCTCAGGAAGTTG 400
Db      125  ArgArgGluGlyArgGluGluGluGluAlaMetGluArg---AlaAlaTyrLeu 143
QY      401  CCTCAACAGTGGCGCCTTAGG---GCACCACAGCGTTGCGACTTGGACGTCGAAAGT 454
Db      144  ProAsnThrCysAsnValArgGluProProArgArgCysAspIleGlnArgHisSer 162

```

Search completed: May 15, 2006, 22:21:58
 Job time : 26.9951 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 21:23:21 ; Search time 32.6702 Seconds
(without alignments)

4645.181 Million cell updates/sec

Title: US-10-728-323-2

Perfect score: 1281

Sequence: 1 gctcaccatactagtagcc.....taaaagatcatgtttgtt 717

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DBV=xlp
-Q=/abs/ABSSWEB_spool/US10728323/runat_15052006_172132_22377/app_query_fasta_1
-DB=UniProt -QFWT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abes02p
-USER=US10728323 @CCN 1.1 694 @runat_15052006_172132_22377 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	842	65.7	179	2	Q647H0 ARAHY
2	838	65.4	175	2	Q6PSU1 ARAHY
3	819	63.9	156	2	Q941R0 ARAHY
4	815	63.6	172	2	Q8GV20 ARAHY
5	815	63.6	187	2	Q6PSU2 ARAHY
6	808	63.1	169	2	Q7Y1C0 ARAHY
7	792	61.8	166	2	Q84TU1 ARAHY
8	454.5	35.5	145	2	Q647G9 ARAHY
9	443.5	34.6	144	2	Q8W251 ARAHY
10	443.5	34.6	145	2	Q516T1 ARAHY
11	401	31.3	124	2	Q511E9 ARAHY
12	388.5	30.3	129	2	Q9SQG5 ARAHY
13	300	23.4	160	2	Q9SQH1 ARAHY
14	299	23.3	158	2	Q647G8 ARAHY
15	278	21.7	153	2	Q99235 LUPAN
16	243	19.0	158	1	2SS_SOYBN

17	243	19.0	158	2	Q53WV6 SOYBN	Q53WV6 glycine max
18	241.5	18.9	155	2	Q9ZN24 SOYBN	Q9ZN24 glycine max
19	187.5	14.6	80	1	CGD2L LUPAN	P09931 lupinus ang
20	184.5	14.4	323	2	Q39928 HELAN	Q39928 helianthus
21	177.5	13.9	168	2	Q8LPD3 LINUS	Q8LPD3 linum usita
22	173.5	13.5	141	1	2SS_CUGMA	Q39649 cucurbita m
23	167	13.0	160	2	Q71HNI_FICAW	Q71HNI ficus awkeo
24	164.5	12.8	148	2	Q9XHP1 SESIN	Q9XHP1 sesamum ind
25	159.5	12.5	143	2	Q8VX62 SESIN	Q8VX62 sesamum ind
26	159	12.4	161	2	Q7Y1C2 JUGNI	Q7Y1C2 juglans nig
27	153.5	12.0	139	2	P93198_GROSI	P93198 juglans reg
28	153	11.9	153	2	Q9AUD1 SESIN	Q9AUD1 sesamum ind
29	151	11.8	295	1	2SSS_HELAN	P15461 helianthus
30	149.5	11.7	167	2	Q84JW2 VITVI	Q84JW2 vitis vinif
31	148.5	11.6	140	2	Q8L694 MOMCH	Q8L694 momordica c
32	148	11.6	285	2	Q8GUD8 HELAN	Q8GUD8 helianthus
33	147	11.5	186	1	2SSE_BRANA	P09893 brassica na
34	146.5	11.4	164	1	2SS1_ARATH	P15457 arabidopsis
35	144.5	11.3	165	2	Q9PH31_ARATH	Q9PH31 arabidopsis
36	144	11.2	169	2	Q8LPD4 LINUS	Q8LPD4 linum usita
37	143.5	11.2	179	2	Q84NG9 VITVI	Q84NG9 vitis vinif
38	141.5	11.0	138	2	Q8H2B8 ANAOC	Q8H2B8 anacardium
39	141	11.0	126	2	Q8FW54_HELAN	Q8FW54 helianthus
40	140.5	11.0	146	1	2SS_BEREX	P04403 bertholleti
41	139	10.9	155	1	2SS2_CAPMA	P30233 capparis ma
42	139	10.9	184	2	Q42444 BRACM	Q42444 brassica ca
43	139	10.9	258	1	2SS_RICCO	P01089 ricinus com
44	132.5	10.3	165	2	Q40850 PICGL	Q40850 picea glauc
45	132.5	10.3	167	2	Q81411_PICGL	Q81411 picea glauc

ALIGNMENTS

RESULT 1

Q647H0 ARAHY
ID Q647H0 ARAHY PRELIMINARY; PRT; 179 AA.
AC Q647H0;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE 2S protein 1 (Fragment).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Yan Y.S., Lin X.D., Zhang Y.S., Wang L., Wu K., Huang S.Z.;
RT "Isolation of peanut genes encoding arachins and conglutins by
RT expressed sequence tags.";
RL Plant Sci. 169:439-445(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Yan Y., Wang L., Huang S.;
RT "cDNA clone of peanut seed storage protein gene.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY722689; AAU21494.1; -; mRNA.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_amyl; 1.
DR SMART; SM00499; AAI; 1.
FT NON_TER 1
SQ SEQUENCE 179 AA; 20850 MW; 68CE6453B489E1DC CRC64;

Alignment Scores:

Pred. No.: 2.99e-65 Length: 179
Score: 842.00 Matches: 157
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 65.7% Indels: 0
DB: 2 Gaps: 0

US-10-728-323-2 (1-717) x Q647H0_ARAHY (1-179)

QY 2 CTCACCATCTAGTACGCTCGCCCTTTCCTCTCGCTGCGCCACCGCATCTCGGAGCGAG 61

Db 23 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaHisAlaSerAlaArgGln 42

QY 62 CAGTGGGAACCTCAAGGAGACAGAAGATGCCAGAGCCAGCTCCAGAGGGCGAACCTTGAGG 121

Db 43 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 62

QY 122 CCCTGGCAGCAACATCTCATGCAAGAAGATCCAAAGTCAGCAGGATTCATATGAACGGGAC 181

Db 63 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspSerTyrGluArgAsp 82

QY 182 CCTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGCGCTGGA 241

Db 83 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 102

QY 242 TCCTCTCAGCACCAGAGAGGTGCTCAATGAGCTGAACGAGTTTGAGACACCAAGG 301

Db 103 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 122

QY 302 TGCATGTGCGAGGATTCGCAACAGATCATGGAGAACCGAGCGCATAGGTTGCAAGGAGG 361

Db 123 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 142

QY 362 CAACAGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCCCTCAACAGTCGCGCTTAGG 421

Db 143 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 162

QY 422 GCACACAGCGTTGCGACTTGCGAGTCGAAAGTGGCGGCAGACAGATAC 472

Db 163 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 179

RESULT 2

Q6PSU1_ARAHY

ID Q6PSU1_ARAHY PRELIMINARY; PRT; 175 AA.

AC Q6PSU1; 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Seed storage protein SSP2 (Fragment).

OS Arachis hypogaea (Peanut).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;

OC Arachis.

OX NCBI_TaxID=3818;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Yan Y., Wang L., Liao B., Li H., Lin X., Huang S.;

RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY581854; AAT00599.1; -; mRNA.

DR InterPro; IPR003612; AAI.

DR Pfam; PF00234; Tryp_alpha_aml; 1.

DR SMART; SM00499; AAI; 1.

FT NON TER 1

SQ SEQUENCE 175 AA; 20463 MW; B3FB5806B70733DE CRC64;

Alignment Scores:

Pred. No.: 6.7e-65 Length: 175

Score: 838.00 Matches: 156

Percent Similarity: 99.4% Conservative: 0

Best Local Similarity: 99.4% Mismatches: 1

Query Match: 65.4% Indels: 0

DB: 2 Gaps: 0

US-10-728-323-2 (1-717) x Q6PSU1_ARAHY (1-175)

QY 2 CTCACCATCTAGTACGCTCGCCCTTTCCTCTCGCTGCGCCACCGCATCTCGGAGCGAG 61

Db 19 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaHisAlaSerAlaArgGln 38

QY 62 CAGTGGGAACCTCAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTGAGG 121

Db 39 GlnTrpGluPheGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 58

QY 122 CCCTCGGAGCAACATCTCATGCAAGAAGATCCAACGTGACGAGGATTCATATGAACGGGAC 181

Db 59 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 78

QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 241

Db 79 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 98

QY 242 TCCTCTCAGCACCAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGACACCAAGG 301

Db 99 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 118

QY 302 TGCATGTGCGAGGATTCGCAACAGATCATGGAGAACCGAGCGCATAGGTTGCAAGGAGG 361

Db 119 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 138

QY 362 CAACAGGAGCAACAGTTCAAGAGGAGGCTCAGGAACCTTGCCCTCAACAGTCGCGCTTAGG 421

Db 139 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 158

QY 422 GCACACAGCGTTGCGACTTGCGAGTCGAAAGTGGCGGCAGACAGATAC 472

Db 159 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 175

RESULT 3

Q941R0_ARAHY

ID Q941R0_ARAHY PRELIMINARY; PRT; 156 AA.

AC Q941R0; 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Allergen II (Fragment).

OS Arachis hypogaea (Peanut).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;

OC Arachis.

OX NCBI_TaxID=3818;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE-21192158; PubMed-11295663; DOI=10.1067/mai.2001.113522;

RA Viquez O.M., Summer C.G., Dodo H.W.;

RT "Isolation and molecular characterization of the first genomic clone

of a major peanut allergen, Ara h 2.,"

J. Allergy Clin. Immunol. 107:713-717(2001).

DR EMBL; AY007229; AAK96887.1; -; Genomic_DNA.

DR InterPro; IPR003612; AAI.

DR Pfam; PF00234; Tryp_alpha_aml; 1.

DR SMART; SM00499; AAI; 1.

FT NON TER 156

SQ SEQUENCE 156 AA; 18167 MW; 664D6832F91F36D1 CRC64;

Alignment Scores:

Pred. No.: 3.1e-63 Length: 156

Score: 819.00 Matches: 153

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 63.9% Indels: 0

DB: 2 Gaps: 0

US-10-728-323-2 (1-717) x Q941R0_ARAHY (1-156)

QY 2 CTCACCATCTAGTACGCTCGCCCTTTCCTCTCGCTGCGCCACCGCATCTCGGAGCGAG 61

Db 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaHisAlaSerAlaArgGln 23

QY 62 CAGTGGGAACCTCAAGGAGAGCAGAAGATGCCAGAGCGAGCTCGAGAGGGCGAACCTGAGG 121

Db 24 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43


```

Db      59 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGlyArgAsp 78
      |||
QY      182 CGGTACAGCCCTAGTCAG-----GATCCG 205
      |||
Db      79 ProTyrSerProSerGlnAspProTyrSerProSerGlnAspProAspArgAspPro 98
      |||
QY      206 TACAGCCCTAGTCCATATGATCGGAGAGCGCTGGATCCTCTCAGCACCAAGAGAGGTGT 265
      |||
Db      99 TyrSerProSerProTyrAspArgGlyAlaGlySerSerGlnHisGlnGluArgCys 118
      |||
QY      266 TGCATAGAGCTGAACAGATTGAGAACCAACCAAGGTGCATGTCGAGGCATTGCAACAG 325
      |||
Db      119 CysAsnGluLeuAsnGluPheGluAsnAsnGlnArgCysMetCysGluAlaLeuGlnGln 138
      |||
QY      326 ATCATGAGAACACAGAGCGATAGTTGCAGGGGAGGCAACAGGAGCAACAGTTCAAGAGG 385
      |||
Db      139 IleMetGluAsnGlnSerAspArgLeuGlnGlyArgGlnGlnGlnGlnPheLysArg 158
      |||
QY      386 GAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGGGCCACACAGGCTTGGCAGCTTGGAC 445
      |||
Db      159 GluLeuArgAsnLeuProGlnGlnCysGlyLeuArgAlaProGlnArgCysAspLeuGlu 178
      |||
QY      446 GTCGAAAGTGGCGCGAGACAGATAC 472
      |||
Db      179 ValGluSerGlyGlyArgAspArgTyr 187
      |||

```

RESULT 6

```

Q7Y1C0_ARAHY
ID   Q7Y1C0_ARAHY PRELIMINARY;          PRT;   169 AA.
AC   QY1C0;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DI   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DR   Allexgen Ara h 2 isoform (Fragment).
OS   Arachis hypogaea (Peanut).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC   Arachis.
OX   NCBI_TaxID=3818;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Becker W.-M., Suhr M., Lindner B., Wicklein D., Lepp U.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR   EMBL; AY117434; AAM78596.1; -; mRNA.
DR   InterPro; IPR003612; AAI.
DR   Pfam; PF00234; Tryp_alpha_amyl; 1.
DR   SMART; SM00499; AAI; 1.
FT   NON_TER 1
SQ   SEQUENCE 169 AA; 19768 MW; 0B14A7ED911F34EF CRC64;

```

```

Alignment Scores:
Pred. No.:      2,89e-62      Length:      169
Score:          808.00        Matches:    154
Percent Similarity: 91.7%      Conservative: 1
Best Local Similarity: 91.1%    Mismatches: 2
Query Match:     63.1%         Indels:    12
DB:              2            Gaps:      1

```

US-10-728-323-2 (1-717) x Q7Y1C0_ARAHY (1-169)

```

QY      2 CTCACCATAGTAGAGCCCTCGCCCTTTCTCCTCGCTGCCACCGCATCTGGAGGCAG 61
      |||
Db      1 LeuThrIleLeuValAlaProAlaLeuPheLeuAlaAlaHisAlaSerAlaArgGln 20
      |||
QY      62 CAGTGGGAACCTCCAAGAGACAGAGATGCCAGAGCCAGCTCCGAGAGGGCGAACCTCAGG 121
      |||
Db      21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
      |||
QY      122 CCCTGCCAGCAACATCTCATGCAGAAATCCAAAGTCAGCAGGATTCATATGAACGGGAC 181
      |||
Db      41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGlyArgAsp 60
      |||

```

```

QY      182 CGGTACAGCCCTAGTCAG-----GATCCG 205
      |||
Db      61 ProTyrSerProSerGlnAspProTyrSerProSerGlnAspProAspArgArgAspPro 80
      |||
QY      206 TACAGCCCTAGTCCATATGATCGGAGAGCGCTGGATCCTCTCAGCACCAAGAGAGGTGT 265
      |||
Db      81 TyrSerProSerProTyrAspArgGlyAlaGlySerSerGlnHisGlnGluArgCys 100
      |||
QY      266 TGCATAGAGCTGAACAGATTGAGAACCAACCAAGGTGCATGTCGAGGCATTGCAACAG 325
      |||
Db      101 CysAsnGluLeuAsnGluPheGluAsnAsnGlnArgCysMetCysGluAlaLeuGlnGln 120
      |||
QY      326 ATCATGAGAACACAGAGCGATAGTTGCAGGGGAGGCAACAGGAGCAACAGTTCAAGAGG 385
      |||
Db      121 IleMetGluAsnGlnSerAspArgLeuGlnGlyArgGlnGlnGlnGlnPheLysArg 140
      |||
QY      386 GAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGGGCCACACAGGCTTGGCAGCTTGGAC 445
      |||
Db      141 GluLeuArgAsnLeuProGlnGlnCysGlyLeuArgAlaProGlnArgCysAspLeuGlu 160
      |||
QY      446 GTCGAAAGTGGCGCGAGACAGATAC 472
      |||
Db      161 ValGluSerGlyGlyArgAspArgTyr 169
      |||

```

RESULT 7

```

Q84TU1_ARAHY
ID   Q84TU1_ARAHY PRELIMINARY;          PRT;   166 AA.
AC   Q84TU1;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DI   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DR   Seed storage protein (Fragment).
OS   Arachis hypogaea (Peanut).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC   Arachis.
OX   NCBI_TaxID=3818;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Paik-Ro O.G., Seib J.C., Smith R.L.;
RL   PubMed=12582692;
DR   EMBL; AF366560; AAC61750.1; -; mRNA.
DR   InterPro; IPR003612; AAI.
DR   Pfam; PF00234; Tryp_alpha_amyl; 1.
DR   SMART; SM00499; AAI; 1.
FT   NON_TER 1
SQ   SEQUENCE 166 AA; 19321 MW; BD04F7F26CE7B437 CRC64;

```

```

Alignment Scores:
Pred. No.:      7,35e-61      Length:      166
Score:          792.00        Matches:    151
Percent Similarity: 92.1%      Conservative: 1
Best Local Similarity: 91.5%    Mismatches: 1
Query Match:     61.8%         Indels:    12
DB:              2            Gaps:      1

```

US-10-728-323-2 (1-717) x Q84TU1_ARAHY (1-166)

```

QY      2 CTCACCATAGTAGAGCCCTCGCCCTTTCTCCTCGCTGCCACCGCATCTGGAGGCAG 61
      |||
Db      2 LeuThrIleLeuValAlaLeuAlaLeuPheLeuAlaAlaHisAlaSerAlaArgGln 21
      |||
QY      62 CAGTGGGAACCTCCAAGAGACAGAGATGCCAGAGCCAGCTCCGAGAGGGCGAACCTCAGG 121
      |||
Db      22 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 41
      |||
QY      122 CCCTGCCAGCAACATCTCATGCAGAAATCCAAAGTCAGCAGGATTCATATGAACGGGAC 181
      |||

```

```

Db      42  ProCysGluGlnHisLeuMetGlnYsIleGlnArgAspGluAspSerTyrGlyArgAsp 61
QY      182  CGGTACAGCCCTAGTCAG-----GATCCG 205
Db      62  ProTyrSerProSerGlnAspProTyrSerProSerGlnAspProAspArgAspPro 81
QY      206  TACAGCCCTAGTCATATGATCGGAGAGCGGTGGATCCTCTCAGACCAAGAGAGGTGT 265
Db      82  TyrSerProSerProTyrAspArgArgGlyAlaGlySerGlnHisGlnGluArgCys 101
QY      266  TGCATGAGCTGAACAGATTGAGAACCAACCAAGAGTGTGTCGAGGCGATTCGAACAG 325
Db      102  CysAsnGluLeuAsnGluPheGluAsnAsnGlnArgCysMetCysGluAlaLeuGlnGln 121
QY      326  ATCATGAGCAACACAGAGCATAGTTGCAGGGGAGGCAACAGGACCAACAGTTCAAGAGG 385
Db      122  IleMetGluAsnGlnSerAspArgLeuGlnGlyArgGlnGlnGlnGlnPheYsArg 141
QY      386  GAGCTCAGGAACCTGCTCAACAGATGCGGCGCTTAGGGCAACACAGCGTTGCGACTTGGAC 445
Db      142  GluLeuArgAsnLeuProGlnGlnCysGlyLeuArgAlaProGlnArgCysAspLeuGlu 161
QY      446  GTCGAAAGTGGCGGC 460
Db      162  ValGluSerGlyGly 166

RESULT 8
Q647G9_ARAHY
ID      Q647G9_ARAHY PRELIMINARY; PRT; 145 AA.
AC      Q647G9;
DT      25-OCT-2004 (TrEMBLrel. 28, Created)
DT      25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT      13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE      Conglutinin.
OS      Arachis hypogaea (Peanut).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC      rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC      Arachis.
OX      NCBI_TaxID=3818;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RA      Yan Y., Wang L., Huang S.;
RT      "cDNA clone of peanut seed storage protein gene.";
RL      Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      NUCLEOTIDE SEQUENCE.
RA      Yan Y., Wang L., Huang S.;
RL      Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AY722690; AAU21495.1; -; mRNA.
DR      EMBL; AY849314; AAW32558.1; -; mRNA.
DR      InterPro; IPR003612; AAI.
DR      Pfam; PF00234; Tryp_alpha_aml1; 1.
DR      SMART; SM00499; AAI; 1.
SQ      SEQUENCE 145 AA; 16920 MW; FA3D6FC41CB6267D CRC64;

```

```

Alignment Scores:
Pred. No.:      3,31e-31      Length:      145
Score:          454.50      Matches:      93
Percent Similarity: 69.7%      Conservative: 15
Best Local Similarity: 60.0%      Mismatches: 30
Query Match:      35.5%      Indels:      17
DB:              2          Gaps:      4

```

US-10-728-323-2 (1-717) x Q647G9_ARAHY (1-145)

```

QY      5  ACCATACTAGTAGCCCTCGCCCTTTCTCTCGTCGCCACGCACTCTGCG---AGGCAG 61
Db      5  ThrIleLeuValAlaLeuLeuAlaLeuValLeuValAlaHisAlaSerAlaMetArgArg 24
QY      62  CAGTGGGAACCTCCAAGGAGAC---AGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTG 118
Db      25  GluArgGlyArgGlnGlyAspSerSerCysGluArgGlnValAspArgValAsnLeu 44
QY      119  AGGCCCTCGAGCAACATCTCATGCAGAAAGATCCAACTGTGACGAGGATTCATATGAACGG 178
Db      45  LysProCysGluGlnHisIleMetGlnArgIleMetGlyGluGlnGlnGlnTyrAsp--- 63
QY      179  GACCCGTACAGCCCTAGTCAGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCT 238
Db      64  -----SerTyrAspIleArgSerThr 70
QY      239  GGATCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGACGAGTTTGAGAACCAACAA 298
Db      71  ArgSerSerAspGlnGlnArgCysCysAspGluLeuAsnGluMetGluAsnThrGln 90
QY      299  AGGTGCATGTGCGAGGCATTGCAACAGATCATGGAGAACCAAGAGCGATAGGTTGCAGGGG 358
Db      91  ArgCysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnCysAspArgLeuGlnAsp 110
QY      359  AGGCAACAGAGGACACAGTTTCAAGAGGAGGAGCTCAGAGAACTTGCTCTCAACAGTGC 418
Db      111  ArgGlnMetValGlnGlnPheYsArgGluLeuMetAsnLeuProGlnGlnCysAsnPhe 130
QY      419  AGGGCACACAGCGTTGCGACTTGAGCTCGAAAGTGGCGGGCAGA 463
Db      131  ArgAlaProGlnArgCysAspLeuAspVal---SerGlyGlyArg 144

RESULT 9
Q8W251_ARAHY
ID      Q8W251_ARAHY PRELIMINARY; PRT; 144 AA.
AC      Q8W251;
DT      01-MAR-2002 (TrEMBLrel. 20, Created)
DT      01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT      01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE      Conglutinin (Fragment).
OS      Arachis hypogaea (Peanut).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC      eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC      Arachis.
OX      NCBI_TaxID=3818;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RA      Paik-Ro O.G., Seib J.C., Smith R.L.;
RT      "Seed-specific, developmentally regulated genes of peanut.";
RL      Theor. Appl. Genet. 104:236-240(2002).
DR      EMBL; AF366561; AAL37561.1; -; mRNA.
DR      InterPro; IPR003612; AAI.
DR      Pfam; PF00234; Tryp_alpha_aml1; 1.
DR      SMART; SM00499; AAI; 1.
FT      NON_TER 1
SQ      SEQUENCE 144 AA; 16750 MW; B776486138A852C3 CRC64;

```

```

Alignment Scores:
Pred. No.:      3.07e-30      Length:      144
Score:          443.50      Matches:      90
Percent Similarity: 69.7%      Conservative: 18
Best Local Similarity: 58.1%      Mismatches: 30
Query Match:      34.6%      Indels:      17
DB:              2          Gaps:      4

```

US-10-728-323-2 (1-717) x Q8W251_ARAHY (1-144)

```

QY      5  ACCATACTAGTAGCCCTCGCCCTTTCTCTCGTCGCCACGCACTCTGCG---AGGCAG 61
Db      4  ThrIleLeuValAlaLeuLeuAlaLeuValLeuValAlaHisAlaSerAlaMetArgArg 23

```

```

QY 62 CAGTGGGAACCTCAAGGAGAC---AGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTG 118
DB 24 GluArgGlyArgGlnGlyAspSerSerCysGluArgGlnValAspArgValAsnLeu 43
QY 119 AGGCCCTCGCAGCAACATCTCATGCAGAGATCCAAACGTGACGAGGATTCATATGAACGG 178
DB 44 LysProCysGluGlnHisIleMetGlnArgIleMetGlyGluGlnGlnTyrAsp--- 62
QY 179 GACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCT 238
DB 63 -----SerTyrAspIleArgSerThr 69
QY 239 GGATCTCTCAGACCAAGAGAGGTGTGCAATGAGCTGAACAGTTCGAGAACACCAA 298
DB 70 ArgSerSerAspGlnGlnArgCysAspGluLeuAspGlnMetGluAsnThrGlu 89
QY 299 AGGTGATGTCGAGGATTCGCAACAGATCATCGAGAACCCAGAGCGATGTTGCAGGGG 358
DB 90 ArgCysMetCysGluAlaLeuGlnIleMetGluAsnGlnCysAspArgLeuGlnAsp 109
QY 359 AGGCAACAGAGCAACAGTTCGAAGAGGAGCTCAGGAACCTTGCTCAACAGTCGCGCCTT 418
DB 110 ArgGlnMetValGlnGlnPheLysArgGluLeuMetAsnLeuProGlnGlnCysAsnPhe 129
QY 419 AGGGCACCAAGGTTGCGACTTGGAGCTCGAAAGTGGCGGCAG 463
DB 130 ArgAlaProGlnArgCysAspLeuAspVal---SerGlyGlyArg 143

```

RESULT 10

```

Q516T1_ARAHY
ID Q516T1_ARAHY PRELIMINARY; PRT; 145 AA.
AC Q516T1;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Conglutin..
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Yan Y., Lin X.D., Zhang Y.S., Wang L., Wu K., Huang S.Z.;
RT "Isolation of peanut genes encoding arachins and conglutins by
RT expressed sequence tags.";
RL Plant Sci. 169:439-445(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Yan Y., Wang L., Huang S.;
RT "cDNA clone of peanut seed storage protein gene.";
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY848699; AAW56068.1; -; mRNA.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_amy1; 1.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 145 AA; 16921 MW; AD139BD008A82C72 CRC64;

```

```

Alignment Scores:
Pred. No.: 3.07e-30 Length: 145
Score: 443.50 Matches: 90
Percent Similarity: 69.7% Conservative: 18
Best Local Similarity: 58.1% Mismatches: 30
Query Match: 34.6% Indels: 17
DB: 2 Gaps: 4

```

US-10-728-323-2 (1-717) x Q516T1_ARAHY (1-145)

```

QY 5 ACCATACTAGTCCCTCGCCCTTTCTCTCTCGTCCCAACGCACTGCG---AGGCAG 61
DB 5 ThrIleLeuValAlaLeuLeuAlaLeuValLeuValAlaHisAlaSerAlaMetArgArg 24

```

```

QY 62 CAGTGGGAACCTCAAGGAGAC---AGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTG 118
DB 25 GluArgGlyArgGlnGlyAspSerSerCysGluArgGlnValAspArgValAsnLeu 44
QY 119 AGGCCCTCGCAGCAACATCTCATGCAGAGATCCAAACGTGACGAGGATTCATATGAACGG 178
DB 45 LysProCysGluGlnHisIleMetGlnArgIleMetGlyGluGlnGlnTyrAsp--- 63
QY 179 GACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCT 238
DB 64 -----SerTyrAspIleArgSerThr 70
QY 239 GGATCTCTCAGACCAAGAGAGGTGTGCAATGAGCTGAACAGTTCGAGAACACCAA 298
DB 71 ArgSerSerAspGlnGlnArgCysAspGluLeuAspGlnMetGluAsnThrGlu 90
QY 299 AGGTGATGTCGAGGATTCGCAACAGATCATCGAGAACCCAGAGCGATGTTGCAGGGG 358
DB 91 ArgCysMetCysGluAlaLeuGlnIleMetGluAsnGlnCysAspArgLeuGlnAsp 110
QY 359 AGGCAACAGAGCAACAGTTCGAAGAGGAGCTCAGGAACCTTGCTCAACAGTCGCGCCTT 418
DB 111 ArgGlnMetValGlnGlnPheLysArgGluLeuMetAsnLeuProGlnGlnCysAsnPhe 130
QY 419 AGGGCACCAAGGTTGCGACTTGGAGCTCGAAAGTGGCGGCAG 463
DB 131 ArgAlaProGlnArgCysAspLeuAspVal---SerGlyGlyArg 144

```

RESULT 11

```

Q511E9_ARAHY
ID Q511E9_ARAHY PRELIMINARY; PRT; 124 AA.
AC Q511E9;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Ara h 6 allergen (Peanut).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Schocker F., Sunr M., Becker W.M.;
RT "Epitope mapping of Ara h 6.";
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY871100; AAW34231.1; -; Genomic_DNA.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_amy1; 1.
DR SMART; SM00499; AAI; 1.
FT NON_TER 1
SQ SEQUENCE 124 AA; 14719 MW; 34F54A8FC7962189 CRC64;

```

```

Alignment Scores:
Pred. No.: 1.64e-26 Length: 124
Score: 401.00 Matches: 79
Percent Similarity: 68.6% Conservative: 15
Best Local Similarity: 57.7% Mismatches: 27
Query Match: 31.3% Indels: 16
DB: 2 Gaps: 3

```

US-10-728-323-2 (1-717) x Q511E9_ARAHY (1-124)

```

QY 56 AGGCAGAGGTGGAACTCCAAAGGAGAC---AGAAGATGCCAGAGCCAGCTCGAGAGGGCG 112
DB 2 ArgArgGluArgGlyArgGlnGlyAspSerSerCysGluArgGlnValAspArgVal 21
QY 113 AACCTGAGGCCCTCGCAGCAACATCTCATGCAGAGATCCCAACGTGACGAGGATTCATAT 172
DB 22 AsnLeuLysProCysGluGlnHisIleMetGlnArgIleMetGlyGluGlnGlnTyr 41
QY 173 GAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGA 232

```

```
Db 42 Asp-----SerTyrAspIleArg 47
::: ||||| |||
233 GCGCTGGATCCTTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAAC 292
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
48 SerThrArgSerSerAspGlnGlnArgCysCysAspGluLeuAsnGluMetGluAsn 67
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
293 AACCAAGGTGCATGTCGAGGCGATTGCAACAGATCATGGAGAACAGAGCGATAGTTG 352
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
68 ThrGlnGlyCysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnCysAspArgLeu 87
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
353 CAGGGAGGCAACAGAGCAACAGTTCAAGAGGAGCTCAGAACTTGCTCAACGTGC 412
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
88 GlnAspArgGlnMetValGlnGlnPheLysArgGluLeuMetSerLeuProGlnGlnCys 107
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
413 GGCCTTAGGCGCACCAAGCGTTGCGACTTGGAGCTCGAAAGTGGCGGCAGA 463
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
108 AsnPheArgAlaProGlnArgCysAspLeuAspVal---SerGlyGlyArg 123
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
Q9SQG5_ARAHY
ID Q9SQG5_ARAHY PRELIMINARY; PRT; 129 AA.
AC Q9SQG5_ARAHY
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Allergen Arah6 (Fragment).
GN Name=Ara h 6;
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]_TaxID=3818;
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Seed;
RX MEDLINE=99406463; PubMed=10474031; DOI=10.1159/000024203;
RA Kleber-Janke T., Cramer R., Appenzeller U., Schlaak M., Becker W.-M.;
RT "Selective cloning of peanut allergens, including profilin and 2S
RT albumins, by phage display technology.";
RL Int. Arch. Allergy Immunol. 119:265-274 (1999).
DR EMBL; AF092846; AAD56337.1; -; mRNA.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_aml1; 1.
DR SMART; SM00499; AAI; 1.
FT NON_TER
SQ SEQUENCE 129 AA; 14929 MW; 4048D3418BB9EB40 CRC64;

Alignment Scores:
Pred. No.: 2,07e-25 Length: 129
Score: 388.50 Matches: 80
Percent Similarity: 66.4% Conservative: 15
Best Local Similarity: 55.9% Mismatches: 31
Query Match: 30.3% Indels: 17
DB: 2 Gaps: 4

US-10-728-323-2 (1-717) x Q9SQG5_ARAHY (1-129)
QY 41 GCCCAGCATCTGCG---AGGCAGCATGGGAACCTCAAGAGAC---AGAAGATGCCAG 94
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 AlaHieAlaSerAlaMetArgArgGluArgGlyArgGlnGlyAspSerSerCysGlu 20
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 95 AGCCAGCTCGAGAGGCGCACTGAGGCCCTCGGACACATCTCATGCAAGATCCAA 154
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21 ArgGlnValAspGlyValAsnLeuLysProCysGluGlnHieIleMetGlnArgIleMet 40
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 155 CBTGACGAGGATTTCATATGAACGGACCGCTACAGCCCTAGTCAGGATCCGTACGCCCT 214
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 41 GlyGluGlnGluGlnTyrAsp----- 47
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 215 AGTCCATATGATCGGAGGCGCTGGATCCTCTCAGCAACCAAGAGAGGTGTGCAATGAG 274
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 48 ---SerTyrAsnPheGlySerThrArgSerSerAspGlnGlnGlnArgCysCysAspGlu 66
275 CTGAACGAGTTTGGAGAACAAACAAAGGTGCGATGTGCGAGGCAATTGCAACAGATCATGGAG 334
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 67 LeuAsnGluMetGluAsnThrGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGlu 86
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 335 AACCAAGCGATAGTTGCGAGGAGGCAACAGGAGCAACAGTTCAAGAGGAGCTCAGG 394
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 87 AsnGlnCysAspGlyLeuGlnAspArgGlnMetValGlnHiePheLysArgGluLeuMet 106
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 395 RACTTTCCTCAACAGCTGCGGCTTAGGCACACACAGCTTGCAGCTTGCACCTCGAAGCT 454
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 107 AsnLeuProGlnGlnCysAsnPheGlyAlaProGlnArgCysAspLeuAspVal---Ser 125
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 455 GCGCGCAGA 463
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126 GlyGlyArg 128
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
Q9SQH1_ARAHY
ID Q9SQH1_ARAHY PRELIMINARY; PRT; 160 AA.
AC Q9SQH1_ARAHY
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Allergen.
GN Name=Ara h 7;
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]_TaxID=3818;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Virginia; TISSUE=Seed;
RX MEDLINE=99406463; PubMed=10474031; DOI=10.1159/000024203;
RA Kleber-Janke T., Cramer R., Appenzeller U., Schlaak M., Becker W.-M.;
RT "Selective cloning of peanut allergens, including profilin and 2S
RT albumins, by phage display technology.";
RL Int. Arch. Allergy Immunol. 119:265-274 (1999).
DR EMBL; AF091737; AAD56719.1; -; mRNA.
DR InterPro; IPR003612; AAI.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 160 AA; 18417 MW; 9F9E4CEE6808D4C CRC64;

Alignment Scores:
Pred. No.: 1.26e-17 Length: 160
Score: 300.00 Matches: 72
Percent Similarity: 61.8% Conservative: 30
Best Local Similarity: 43.6% Mismatches: 33
Query Match: 23.4% Indels: 31
DB: 2 Gaps: 7

US-10-728-323-2 (1-717) x Q9SQH1_ARAHY (1-160)
QY 2 CTCACCATAGTAGCCCTC-----GCCCTTTTCTCTCTCGCTGCC----- 43
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 5 LeuSerIleLeuValAlaLeuLeuGlyAlaLeuLeuValAlaSerAlaThrArgTyr 24
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 44 -----CACGCATCTCGAGGCGAGCGAGCTGGGAACCTC-----CAAGGAGACAGA 88
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25 AspProAspArgGlySerArgGlySerArgTyrAspAlaProSerArgGlyAspGln 44
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 89 TCCAGAGCCAGCTCGAGAGGCGAGCCTGAGGCCCTCGGACCAACATCTCATGCAAGAG 148
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 45 CysGlnArgGlnLeuGlnArgAlaAsnLeuArgProCysGluGluHieMetArgArg 64
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 149 ATCCAACTGACGAGGATTTCATGAACGGACCGCTACAGCCCTAGTCAGGATCCGTAC 208
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 65 ValGluGlnGluGlnGluGln---GluGlnAspGluTyr----- 76
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 209 AGCCCTAGTCCATATGATCGGAGGCGCTGGATCCTCTCTCAG----- 250
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```
QY 206 TACAGCCCTAGTCCATATGATCGAGAGCGCTGGATCCTCTCAGCACCAAGAGAGGTGT 265
Db |||||
80 TyrSerGlu-----GluSerGluGluLeuAspGlnCys 90
QY 266 TGCATGAGCTGAACGAGTTTGAGAACACCAAGGTGCATGTGCGAGGCATTGCAACAG 325
Db ||| :|||
91 CysGluGlnLeuAsnGluLeu---AsnSerGlnArgCysGlnCysArgAlaLeuGlnGln 109
QY 326 ATCATGGAGAACCAAGCGATAGTTTGCAGGGGAGGCAACAGGAGCAACAGTTCAAGAGG 385
Db ||| :|||
110 IleTyrGluSerGlnSerGluGlnCysGluGlySerGlnGlnGluGlnLeuGluGln 129
QY 386 GAGCTCAGGAACCTGCTCAACAGTCGGCCTTAGGGCACCAAGCGTTTCGACTTGGAC 445
Db |||||
130 GluLeuGluLysLeuProArgThrCysGlyPheGlyProLeuArgCysAspValAsn 149
```

Search completed: May 15, 2006, 22:17:08
Job time : 167.351 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 21:23:02 ; Search time 45.6522 Seconds
(without alignments)
4400.310 Million cell updates/sec

Title: US-10-728-323-3

Perfect score: 2779

Sequence: 1 cggcgcaaccggaggagaa.....ctccgaggctgtggtcttaa 1524

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xl

-Q=/abss/ABSSWEB_spool/US10728323/runat_15052006_172130_22366/app_query.fasta.1
-DB=A_Geneseq -QMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi -ALIGN=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss02p
-USER=US10728323 @CGN 1.1 605 @runat_15052006_172130_22366 -NCPU=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A_Geneseq.21.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1980s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*
- 9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2673	96.2	510	6	ABU52484 Peanut Ar
2	2673	96.2	510	7	ADG27544 Peanut al
3	2667	96.0	510	4	Aau04708 Anaphylac
4	2665	95.9	507	8	ADM12139 Arachis h
5	2663	95.8	510	2	Aay15246 Peanut al
6	2652	95.4	512	2	Aay40912 Ara h 3 a
7	2638	94.9	526	3	Aab33601 Modified
8	2638	94.9	526	4	Aau04711 Modified
9	2638	94.9	526	4	Aau05036 Modified

10	2615	94.1	507	8	ADO38357	Ado38357	Peanut al
11	1502	54.0	481	5	ABG71266	Abg71266	Glycine m
12	1502	54.0	481	7	ADH89253	Adh89253	G. max gl
13	1502	54.0	481	7	ADL90187	Adl90187	Soybean g
14	1502	54.0	481	8	ADG43988	Adg43988	G. max gl
15	1502	54.0	488	8	ADG74499	Adg74499	Plant ful
16	1502	54.0	491	8	ADX74681	Adx74681	Plant ful
17	1502	54.0	492	8	ADX74470	Adx74470	Plant ful
18	1502	54.0	492	8	ADX76928	Adx76928	Plant ful
19	1499.5	54.0	466	8	ADY24175	Ady24175	Plant ful
20	1499.5	54.0	466	8	ADY24189	Ady24189	Plant ful
21	1499.5	54.0	485	5	ABG71265	Abg71265	Glycine m
22	1499.5	54.0	485	7	ADH89247	Adh89247	G. max gl
23	1499.5	54.0	485	7	ADL90186	Adl90186	Soybean g
24	1499.5	54.0	485	8	ADG43982	Adg43982	G. max gl
25	1499.5	54.0	488	8	ADX76279	Adx76279	Plant ful
26	1499.5	54.0	489	8	ADY24211	Ady24211	Plant ful
27	1499.5	54.0	489	8	ADY24504	Ady24504	Plant ful
28	1499.5	54.0	489	8	ADY76246	Ady76246	Plant ful
29	1499.5	54.0	489	8	ADY24125	Ady24125	Plant ful
30	1499.5	54.0	489	8	ADX75310	Adx75310	Plant ful
31	1499.5	54.0	489	8	ADY24174	Ady24174	Plant ful
32	1499.5	54.0	489	8	ADX87453	Adx87453	Plant ful
33	1499.5	54.0	489	8	ADY24183	Ady24183	Plant ful
34	1499.5	54.0	489	8	ADX88748	Adx88748	Plant ful
35	1499.5	54.0	489	8	ADX87543	Adx87543	Plant ful
36	1499.5	54.0	489	8	ADY24118	Ady24118	Plant ful
37	1499.5	54.0	489	8	ADY24503	Ady24503	Plant ful
38	1499.5	54.0	489	8	ADY74617	Ady74617	Plant ful
39	1499.5	54.0	489	8	ADY24173	Ady24173	Plant ful
40	1499.5	54.0	489	8	ADX95765	Adx95765	Plant ful
41	1498	53.9	489	8	ADX76273	Adx76273	Plant ful
42	1496	53.8	476	9	ADZ47094	Adz47094	Soybean g
43	1496	53.8	477	9	ADZ47108	Adz47108	Soybean m
44	1496	53.8	485	9	ADZ47102	Adz47102	N-termina
45	1496	53.8	485	9	ADZ47106	Adz47106	C-termina

ALIGNMENTS

RESULT 1

ABUS2484

ID ABUS2484 standard; protein; 510 AA.

XX AC ABUS2484;

XX DT 10-MAR-2003 (first entry)

XX DE Peanut Ara h3 protein sequence.

XX KW Peanut; allergy; Ara h1; Ara h2; Ara h3; IgE binding site;

XX OS Arachis hypogaea.

XX PN WO200274250-A2.

XX PD 26-SEP-2002.

XX PF 18-MAR-2002; 2002WO-US009108.

XX PR 16-MAR-2001; 2001US-0276822P.

XX PR 18-MAR-2002; 2002US-00276822.

XX (PANA-) PANACEA PHARM.

XX Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;

XX Compadre CW, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;

XX Rabinjohn PA, Shin DS, Stanley JS;

XX WPI; 2003-018765/01.

XX DR N-PSDB; ABX70612.

XX

PT New modified anaphylactic food allergen, useful for preventing or
PT treating allergic reactions associated with e.g. anaphylactic allergens.
XX
PS
XX
XX
CC The invention relates to a modified anaphylactic food allergen has an
CC amino acid sequence that is substantially identical to that of natural
CC anaphylactic food allergen, except for a cysteine residue that has been
CC modified so that it cannot participate in the disulphide bond. The
CC modification may also comprise mutation of the IgE binding sites to
CC reduce allergenicity. Also included are: (1) a method of making a
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
CC or for causing a site specific mutation in the modified anaphylactic food
CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified
CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a
CC peanut allergen (e.g. Ara h1, h2 or h3)
XX
SQ Sequence 510 AA;

Alignment Scores:

Pred. No.: 4,28e-236 Length: 510
Score: 2673.00 Matches: 507
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 96.2% Indels: 0
DB: 6 Gaps: 0

US-10-728-323-3 (1-1524) x ABU52484 (1-510)

QY	1	CGCAGCAACCGGAGGAGACGGTTCAGGCGCTCAATGGCGAGACCTGAC	60
DB	4	ArgGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	23
QY	61	AATCGCATTAATCAGAGGCGGTACATTGAGACTTGGACCCCAACACACAGGAGTTC	120
DB	24	AsnArgIleGluSerGluGlyGlyTyrlleGluThrTrpAsnProAsnGlnGluPhe	43
QY	121	GAATGCGCGCGTCCGCTCTCTCGCTTAGTCTCTCGCGCAACGCCCTTCGTAGCCT	180
DB	44	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro	63
QY	181	TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGGGATACTTTGGTTG	240
DB	64	PheTyrSerAsnAlaProGlnGluIlePheileGlnGlnGlyArgGlyTyrPheGlyLeu	83
QY	241	ATATTCCCTGGTTGCTCTAGACACTATGAAGAGCTCACACAAAGTTCGTGATCTCAG	300
DB	84	IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgSerGln	103
QY	301	TCCCAAGACACCAAGACGTCCTCCAGGAGAGACCAAGCCCAACAGCAACGAGATAGT	360
DB	104	SerGlnArgProProArgGLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer	123
QY	361	CACGAGAGTGCACCGTTTCGATGAGGTTGATCTCATCCAGCAAGGAGGGGATACTTTGGTTG	420
DB	124	HisGlnIysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla	143
QY	421	TTCTGGCTCTACACGACACGACACTGATGTTGTTGCTGTTCTTCTTACTGACACCAAC	480
DB	144	PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn	163
QY	481	AACAACGACACACAGCTTGATTCAGTTCGCCAGGAGATTCATTTGGCTGGGAACCGGAG	540
DB	164	AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu	183
QY	541	CAAGAGTTCITTAAGGTACCACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT	600

DB	184	GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgArgSerLeuProTyr	203
QY	601	AGCCCATACAGCCCGCAAGTTCAGCTAGACAAAGACGCTGAATTTAGCCCTCGAGGA	660
DB	204	SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly	223
QY	661	CAGCAGCGGAGAGAACGAGCAGGACAAAGAAAGAAAGAAAGAAAGAAAGAAAGTTC	720
DB	224	GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluAsnGluGlyGlyAsnIlePhe	243
QY	721	ACGGCTTCACGCGGAGTTCCTGGAAACAGGCTTCAGGTTGACGACAGACAGATAGTG	780
DB	244	SerGlyPheThrProGlnPheLeuGlnAlaPheGlnValAspAspArgGlnIleVal	263
QY	781	CAAAACCTAAGAGGCGAGACCGAGAGTGAAGAGGAGGAGCCATTGTGACGTGAGGGA	840
DB	264	GlnAsnLeuArgGlyGluThrGluSerGluGluGlyAlaIleValThrValArgGly	283
QY	841	GGCTCAGAAATCTTGAGCCCATAGAAAGACGCTGCCGACGAAAGAGGAAAGATACGAT	900
DB	284	GlyLeuArgIleLeuSerProAspArgLysArgAlaAspGluGluGluGlyArgAsp	303
QY	901	GAAGATGAATATCAATACGATGAAGAGGATAGAAGCGCTGGCAGGGGAAAGCAGGACG	960
DB	304	GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg	323
QY	961	GGGAATGTTATTGAAGAGACGATCTGCACCGCAAGTGTAAAGAAACATTTGTTGAAAC	1020
DB	324	GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn	343
QY	1021	AGATCCCTGACATCTAGAACCTCAGCTGTTCACTCAAACTGCCACGATCTCAAC	1080
DB	344	ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn	363
QY	1081	CTTCTAATTAATAGTGGCTTGACCTAGTCTGTAATATGGAATCTCTACAGGAATGCA	1140
DB	364	LeuLeuIleuArgIleuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla	383
QY	1141	TTGTTTGTGCTCCTACATCAACACCAACCGCACAGCATCATATATCGATTGAGGGGACG	1200
DB	384	LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg	403
QY	1201	GCTCAGCTGCAAGTCTGACAGCAACCGCAACAGAGTGTACGACGAGGAGTTCAGAG	1260
DB	404	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluLeuGlnGlu	423
QY	1261	GGTCACGTGCTTGTGTCACACAACTTCGCGCTCGCTGGAAAGTCCACAGACGAGAAC	1320
DB	424	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn	443
QY	1321	TTCAATATCGTGCATTCAGACAGACTCAAGCCCGCAGATAGCCACCTCGCCGCTGAA	1380
DB	444	PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu	463
QY	1381	AATCCGTCATAGATAACCTGCGGAGAGGTTGTTGCAAAATTCATATGSCCTCCAAAGG	1440
DB	464	AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg	483
QY	1441	GAGCAGGCAAGGCGCTTAAGAACCAACCCCTTCAAGTTCTTCAAGTTCTTCCTCAG	1500
DB	484	GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProProSerGln	503
QY	1501	CAGTCTCCGAGGCTGTGGCT 1521	
DB	504	GlnSerProArgAlaValAla 510	

RESULT 2

ADG27544

ID ADG27544 standard; protein; 510 AA.

XX

AC ADG27544;

XX

181 TTCTACTCCAAATGCTCCCGAGAGATCTTCATCCAGCAAGGAAGGGATCTTTGGTTG 240
 Db PheTyr-SerAsnAlaProGlnGluIlePheIleGlnGlyArgGlyTyrPheGlyLeu 83
 241 ATATTCCCTGGTTGCTTAGACACTATGAAGACCTCACACAAAGTTCGTGATCTCAG 300
 Db IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgSerGln 103
 301 TCCCAAGACCAACCAAGACGCTCTCCAAGGAGAGAGACCAAGCCAAACAGCAACGAGATAGT 360
 Db SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 123
 361 CACCAGAAGGTGCACCGTTTCGATGAGGGTGTCTCATTCAGTCCCAACCGGTGTGCT 420
 Db HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143
 421 TTCTGGCTCTACACACACGACACTGTATGTTGTTGCTGTTCTTCTTACTGACACCAAC 480
 Db PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 163
 481 AACCAACGACAACCCAGCTGTGATCAGTTCCCGAGAGATTCATATTGGCTGGGAACACGGAG 540
 Db AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 183
 541 CAAGAGTCTTAAGGTACACAGCAACCAAGCAGACAAAGCAGACGAAGACTTACCATAT 600
 Db GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 203
 601 AGCCATACAGCCCGCANAGTCAGCTCAGTCAGACAGAGAGCGGTGAATTAGCCCTCAGGA 660
 Db SerProTyr-SerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 223
 661 CACACACGCGCAGAGAACGAGCAGGACCAAGAAAGAAACCAAGGTGGAAACATCTTC 720
 Db GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIlePhe 243
 721 AGCGGCTTACGCGCGAGTTCCTGGAAACAAGCTTCCAGGTTCCACGACACAGATAGTG 780
 Db SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 263
 781 CAAACCTTAAGCGCAGACCGAGAGTGAAGAGAGGGAGCCATTCGTGACGTGAGGGA 840
 Db GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyValAlaIleValThrValArgGly 283
 841 GGCCTCAGATCTTGAGCCAGATAGAAAGACGACGTCGCCAGCAAGAGGAATACGAT 900
 Db GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrAsp 303
 901 GAAGATGAATATGAATACGATGAAGAGATAGAACGGTGGCAGGGGAACGAGCAGG 960
 Db GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyThr 323
 961 GCGAATGGTATTCAAGAGACGATCTGCACCGCAAGTGTAAAGAGACATTGGTAGAAC 1020
 Db GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 343
 1021 AGATCCCTTGACATCTACAACCCCTCAAGCTGGTTCCTCAAACTCCCAACGATCTCAAC 1080
 Db ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 363
 1081 CTTCTAATCTTAGTGGCTGGACCTAGTGTGAATATGGAAATCTCTACAGGAATGCA 1140
 Db LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 383
 1141 TTCTTTGCTCGCTCACTACACCAACCGCACAGCATCATATATCGATTGGGGACCG 1200
 Db LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 403
 1201 GCTCAGGTGCAAGTCTGTGACACCAACCGCAACAGAGTGTACACGAGAGGCTTCAAGAG 1260
 Db AlaHisValGlnValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 423
 1261 GGTACGTGTTGTGTGCCACAGAACTTCGCGCTCGCTGGAAAGTCCCGAGCGGAGAAC 1320

424 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 443
 1321 TTCCGAATACGTGSCATTCGAAGACAGACTCAAGCCCGCAGCATAGCCAACTCGCCGCTGAA 1380
 444 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 463
 1381 AACTCCGTCTATAGATAACCTCGCGAGGAGGTGGTTGCAAAATTCATATATGCGCTCCAAAGG 1440
 464 AsnSerValIleAspAsnLeuProGluGluValAlaAsnSerTyrGlyLeuGlnArg 483
 1441 GAGCAGGCAAGCAGCTTGAAGAACAAACCCCTTCAAGTCTTCGTTCCACCGTCTCAG 1500
 484 GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPhePheValProProSerGln 503
 1501 CAGTCTCCGAGGCGTGTGGCT 1521
 504 GlnSerProArgAlaValAla 510
 RESULT 4
 ADM12139
 ID ADM12139 standard; protein; 507 AA.
 XX ADM12139;
 AC
 XX 20-MAY-2004 (first entry)
 XX Arachis hypogaea 3 (Ara h3) protein.
 XX antigen presentation enhancing hybrid polypeptide; mammalian Ii-Key;
 KW MHC class II; antibacterial; virucide; fungicide; antirheumatic;
 KW antiarthritic; neuroprotective; dermatological; immunosuppressive;
 KW antiinflammatory; antidiabetic; antithyroid; immune;
 KW rheumatoid arthritis; multiple sclerosis; lupus erythematosus;
 KW diabetes mellitus; myasthenia gravis; autoimmune thyroiditis;
 KW scleroderma; dermatomyositis; pemphigus.
 XX Arachis hypogaea.
 OS
 XX US2003235594-A1.
 PN
 XX 25-DEC-2003.
 PD
 XX 17-SEP-2002; 2002US-00245871.
 PF
 XX 14-SEP-1999; 99US-00396813.
 PR
 XX 17-JUL-2002; 2002US-00197000.
 PR
 XX (ANTI-) ANTIGEN EXPRESS INC.
 PA
 XX Humphreys R, Xu M;
 PI
 XX WPI; 2004-070554/07.
 DR
 XX Novel Ii-Key/antigen presentation enhancing hybrid polypeptide, useful
 PT for treating infections, rheumatoid arthritis, multiple sclerosis, lupus
 PT erythematosus and diabetes mellitus.
 XX
 XX Example 3; Page 21; 87pp; English.
 PS
 XX The invention relates to a novel antigen presentation enhancing hybrid
 CC polypeptide. The novel polypeptide has an N-terminal element consisting
 CC of 4-16 residues of a mammalian Ii-Key peptide and its non-N-terminal
 CC deletion modifications, a chemical structure covalently linking the N-
 CC terminal element to an MHC class II-presented epitope of a C-terminal
 CC element. The C-terminal element comprises an antigenic epitope, which
 CC binds to an antigenic peptide binding site of an MHC class II molecule.
 CC The antigen presentation enhancing hybrid polypeptide has the following
 CC activities: antibacterial, virucide, fungicide, antirheumatic,
 CC antiarthritic, neuroprotective, dermatological, immunosuppressive,
 CC antiinflammatory, antidiabetic, and antithyroid. The antigen presentation
 CC enhancing hybrid polypeptide is useful for modulating the immune response
 CC in an individual and for treating infections (such as bacteria, virus,

CC parasite and fungus), rheumatoid arthritis, multiple sclerosis, lupus
CC erythematous, diabetes mellitus, myasthenia gravis, autoimmune
CC thyroiditis, scleroderma, dermatomyositis and pemphigus. This sequence
CC represents a mammalian II key related protein of the invention.
XX
SQ Sequence 507 AA;

Alignment Scores:
Pred. No.: 2,32e-235 Length: 507
Score: 2665.00 Matches: 506
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 95.9% Indels: 0
DB: 8 Gaps: 0

US-10-728-323-3 (1-1524) x ADM12139 (1-507)

QY	1	CGGCAGCAACCGGAGGAGACGGTGCAGTTCAGCGCTCAATGCGCAGACCTGAC	60
Db	1	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	20
QY	61	AATCGCATTAATCAGAGGGCGGTACATTGAGACTTGGAAACCCCAACACCGAGGTTTC	120
Db	21	AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe	40
QY	121	GAATGCGCGCGTCCCTCTCTCGTGTAGTCTCTCGCGCGAAACGCCCTTCGTAGCCT	180
Db	41	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro	60
QY	181	TTCTACTCCAACTGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGATACTTTGGTTG	240
Db	61	PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu	80
QY	241	ATATTCCCTGGTTGTCCTAGACACTATGAAGAGCCCTCACACAAAGTGCTCGATCTCAG	300
Db	81	IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln	100
QY	301	TCCCAAGACCAACAGACGCTCTCCAAGGAGAGACCAAGCCCAACAGCAACGAGATAGT	360
Db	101	SerGlnArgProProArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer	120
QY	361	CACCAGAGGTGCACCGTTTCGATGAGGGTGATCTCATTCAGTTCGCCCGGTGTGCT	420
Db	121	HisGlnLysValHisArgThrAspGluGlyAspLeuIleAlaValProThrGlyValAla	140
QY	421	TTCTGGCTCTCAACCGACACGACTGATGTTGTTGCTGTTCTCTTACTGACCAAC	480
Db	141	PheTrpLeuTyrAsnAspHisAspThrAspValAlaValSerLeuThrAspThrAsn	160
QY	481	AACACGCAACACGCTTGATCAGTTCGCCAGGAGATTCAATTTGGCTGGGAACCGGAG	540
Db	161	AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu	180
QY	541	CAAGAGTTCCTTAAGTACCAGCAACAAAGCAGACAAAGCAGACGAAGCTTACCATAT	600
Db	181	GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgArgSerLeuProTyr	200
QY	601	AGCCCATACAGCCCGCAAGTCAAGCTAGACAAAGAGAGCGTGAATTTACCCCTCGAGGA	660
Db	201	SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly	220
QY	661	CAGCACAGCCGACAGACCGAGCAGGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAG	720
Db	221	GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIlePhe	240
QY	721	AGCGGCTTCACGCGGAGTTTCCTGGAAACAAGCCCTCCAGGTTGACGACAGACAGATAGT	780
Db	241	SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal	260
QY	781	CAAAACCTAAGCGCAGACCGAGAGTGAAGAGAGGAGCCATTGTGACAGTGAAGGGA	840
Db	261	GlnAsnLeuArgGlyGluThrGluSerGluGluGlyAlaIleValThrValArgGly	280

QY	841	GGCCTCAGATCTTGAGCCCGACATAGAAAGACAGCGTGGCGACGACGAGGAATACGAT	900
Db	281	GlyLeuArgIleLeuSerProAspArgLysArgAlaAspGluGluGluGlyArgAsp	300
QY	901	GAAGATGAATATCAATACGATGAAGAGGATAGAAGCGGTGGCAGGGGAACGAGGACAGG	960
Db	301	GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg	320
QY	961	GGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGTGTAATAAAGAACATTTGGTGA	1020
Db	321	GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysAsnIleGlyArgAsn	340
QY	1021	AGTCCCTGACATCTACAAACCTCAAGCTGGTTCACTCAAACTGCCAACCATCTCAAC	1080
Db	341	ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn	360
QY	1081	CTTCTAATATCTTAGTGGCTTGACCTAGTGTCTGAATATGGAATCTCTACAGGAATGCA	1140
Db	361	LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla	380
QY	1141	TTGTTTGTCTGCTACATAACACCAACGACACAGCATCATATATCGATTGAGGGACGG	1200
Db	381	LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg	400
QY	1201	GCTCAGCTGCAAGTCTGTGACAGACGCAACGAGAGTGTACGACGAGGAGCTTCAAGAG	1260
Db	401	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluGlnGlu	420
QY	1261	GGTCACTGCTTGTGTGGTGCACAGAACTTTCGCCGTCTGGAAGTCCACGAGCGAGA	1320
Db	421	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn	440
QY	1321	TTTCAATAGTGGCATTCAGACAGACTCAAGCCCGACAGATAGCCACCTCCGCGGTGA	1380
Db	441	PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu	460
QY	1381	AACTCCGTTCATAGATAACCTCCGAGGAGGTGGTTGCAAAATTCATATGGCCCTCCAA	1440
Db	461	AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg	480
QY	1441	GAGCAGCGAAGGAGGCTTAAGAACAAACCCCTTCAAGTTCCTTCCACCGTCTCAG	1500
Db	481	GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPhePheValProProSerGln	500
QY	1501	CAGTCTCCGAGGCTGTGGCT 1521	
Db	501	GlnSerProArgAlaValAla 507	

RESULT 5

AA15246
ID AA15246 standard; protein; 510 AA.
XX
AC AA15246;
XX
DT 17-OCT-2003 (revised)
DT 09-NOV-1999 (first entry)
XX
DE Peanut allergen, Ara h 3, amino acid sequence.
XX
KW allergy; immune response; transgenic; allergen; epitope;
KW immunoglobulin E; Ig E; binding site; peanut.
XX
OS Arachis hypogaea.
XX
PN WO9938978-A1.
PD
PD 05-AUG-1999.
XX
PF 29-JAN-1999; 99WO-US002031.
XX
PR 31-JAN-1998; 98US-0073283P.
PR 13-FEB-1998; 98US-0074590P.
PR 13-FEB-1998; 98US-0074624P.

PR 13-FEB-1998; 98US-0074633P.
 PR 27-AUG-1998; 98US-00141220.
 XX
 XX (UYAR-) UNIV ARKANSAS.
 PA (UYNY) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.
 PA (SOSI/) UNIV H.
 XX
 XX Sosin H, Bannan GA, Burks AW, Sampson HA;
 DR WPI; 1999-479189/40.
 DR N-PSDB; AA206384.
 XX
 XX Modified allergen with reduced IgE binding, useful for treating e.g.
 PT allergies.
 PT
 XX Disclosure; Page 39-40; 46pp; English.
 XX
 CC This is the amino acid sequence of the Ara h 3 protein from Arachis
 CC hypogaea. The Ara h 3 protein has 4 IGE (Immunoglobulin E) binding
 CC epitopes, one of which is immunodominant (AAV15281). Note: The first
 CC three amino acids of this sequence are not encoded by the related
 CC nucleotide sequence. By modifying the IGE binding sites the ability of
 CC the allergen to provoke an immune response is downregulated. The epitopes
 CC of the IGE binding sites can therefore be modified in genetically
 CC engineered plants and animals to elicit less of an allergic response.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 XX
 XX Sequence 510 AA;
 SQ

Alignment Scores:
 Pred. No.: 3,55e-235 Length: 510
 Score: 2663.00 Matches: 506
 Percent Similarity: 99.8% Conservative: 0
 Best Local Similarity: 99.8% Mismatches: 1
 Query Match: 95.8% Indels: 0
 DB: 2 Gaps: 0

US-10-728-323-3 (1-1524) x AAV15246 (1-510)

QY 1 CGGCACCAACCGAGGAGACGGTCCAGTTCACGCGCTCAATCGCGAGACCTGAC 60
 Db 4 ArgGlnGlnProGluGlnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 23
 QY 61 AATCGCATTAATCAGAGCGCGTTACATTGAGACTTGGAACCCCAACACAGGAGTTTC 120
 Db 24 AsnArgIleGluSerGluGlyGlyTyrlleGluThrTrpAsnProAsnAsnGlnGluPhe 43
 QY 121 GAATGCGCGCGTCCGCTCTCTCGCTTAGTCTCGCTCGCGCGAACCGCCCTTCGTAGGCT 180
 Db 44 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro 63
 QY 181 TTCTACTTCAATGCTCCCGAGGAGATCTTCATCCAGCAGGAGGGGATACTTTGGGTG 240
 Db 64 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 83
 QY 241 ATATTCCCTGGTTGCTCTAGACACTATGAGAGCCTCACACACAGGTGCTGATCTCAG 300
 Db 84 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 103
 QY 301 TCCCAAGACCAACAGAGCTCTCCAGGAGAGACCAAGCCCAACAGCAACAGAGATAGT 360
 Db 104 SerGlnArgProProArgLeuGlnGlyGluAspGlnSerGlnGlnGlnArgAspSer 123
 QY 361 CACCAGAGGTGCACCGTTTCATGAGGTGATCTCATTCAGTTCACCGGTGCTGCT 420
 Db 124 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143
 QY 421 TTCTGGCTCTACACACACACGACTGATGTTGTTGCTGTTCTCTTACTGACACCAAC 480
 Db 144 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 163
 QY 481 AACCAACGACACAGCTTGATCAAGTTCCCGAGGAGATTCAATTTGGCTCGGACACGGAG 540

RESULT 6

AAV40912

ID AAY40912 standard; protein; 512 AA.

Db 164 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 183
 QY 541 CAAGAGTTCTTAAGGTACCAGCAACAAAGACACACAAAGACAGACGAGAGCTTACCATAT 600
 Db 184 GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgSerLeuProTyr 203
 QY 601 AGCCCATACAGCCCGCAAGTCTAGACAAAGAGCGGTGAATTTAGCCCTCGAGGA 660
 Db 204 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgPheSerProArgGly 223
 QY 661 CAGCACGCCCGCAGAACGACGACGACACAGAACAGAACAGAACAGAGTGGAAACATCTTC 720
 Db 224 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluAsnGluGlyGlyAsnIlePhe 243
 QY 721 AGCGGCTTCACGCCGAGTTCTCGAAACAAGCTTCCAGGTTGACGACACAGACAGATAGT 780
 Db 244 SerGlyPheThrProGluPheLeuGlnAlaPheGlnValAspAspArgGlnIleVal 263
 QY 781 CAAAACCTTAAGGCGAGACCGAGAGTGAAGAGAGGGAGCCATTGTGACAGTGGGGA 840
 Db 264 GlnAsnLeuArgGlyGluThrGluSerGluGluGlyAlaIleValThrValArgGly 283
 QY 841 GGCCTCAGNATCTTGGCCCATAGAAAGACACGTCGCGACGAAAGAGGATACGAT 900
 Db 284 GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluGlyTyrAsp 303
 QY 901 GAAGATGAATATGAATACGATGAAGAGGATAGAAGCGTGGCAGGGGAAGACAGAGCAG 960
 Db 304 GluAspGluTyrGluTyrAspGluAspArgArgGlyArgGlySerArgGlyArg 323
 QY 961 GGGAAATGGTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAGAAACATTCGTAGA 1020
 Db 324 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 343
 QY 1021 AGATCCCTGACATCTACACCCCTCAAGCTGTTCACTCAAACTGCCAACCATCTCAAC 1080
 Db 344 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 363
 QY 1081 CTTCCTAACTATTAGTGGCTTCGACCTAGTGTCTGAATATGGAAATCTCTACAGGAATGCA 1140
 Db 364 LeuLeuIleLeuArgTrpLeuGlyLeuSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 383
 QY 1141 TTGTTTGTGCTCTACAAACACCAACCGCACACGATCATATATCGATTGAGGGACGG 1200
 Db 384 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 403
 QY 1201 GCTCAGTCGAGTCTGCGACGACGACCGCACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
 Db 404 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 423
 QY 1261 GGTCACGTGCTTGTGTGTCACAGAACTTCGCGCTCGCTGGAAAGTCCCAGACGCGAGA 1320
 Db 424 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 443
 QY 1321 TTCGAATACGTGGCATTCGAAGACAGACTCAAGCCCGACGATAGCCAACTTCGCGGTGAA 1380
 Db 444 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 463
 QY 1381 AACTCCGTCATAGATAAAGTCCGCGAGGAGGTGGTTCGAATTCATATCGCTCCCAAGG 1440
 Db 464 AsnSerValIleAspAsnLeuProGluValValAlaAlaAsnSerTyrGlyLeuGlnArg 483
 QY 1441 GAGCAGGCAAGGACGCTTAAGAACCAACACCCCTTCAAGTCTTCCTTCCACGCTCTCAG 1500
 Db 484 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProProSerGln 503
 QY 1501 CAGTCTCCGAGGCTGTGCT 1521
 Db 504 GlnSerProArgAlaValAla 510

```

XX AAY40912;
AC
XX
XX 17-OCT-2003 (revised)
DT 06-DEC-1999 (first entry)
XX
XX Ara h 3 allergen sequence.
DE
XX
XX Peanut; allergen; Ara H 1; IgE; immunoglobulin E; epitope; Ara h 3;
KW allergic reaction; ss.
XX
XX Arachis hypogaea.
OS
XX
XX Key Location/Qualifiers
FH 187..188
FT Region
FT
FT /note= "these two amino acid residues are not indicated
FT in the Ara h 3 sequence provided in the sequence listing
FT (seq ID No: 6) and Fig 51 of the specification"
FT Misc-difference 374
FT /note= "encoded by CCT"
XX
XX WO9945961-A1.
PN
XX
XX 16-SEP-1999.
PD
XX
XX 12-MAR-1999; 99WO-US005494.
PF
XX
XX 12-MAR-1998; 98US-0077763P.
PR
XX 11-MAR-1999; 99US-00077763.
PR
XX
XX (UYAR-) UNIV ARKANSAS.
PA
XX
XX Burks W, Helm RM, Cockrell G, Bannon GA, Stanley JS, Shin DS;
PI Sampson H, Compadre CM, Huang SK, Maleki SJ, Kopper RA;
PI
XX
XX WPI; 1999-551218/46.
DR
XX N-PSDB; AAZ22280.
DR
XX
XX Tertiary structure of peanut allergen Ara h 1 for protection of a host
PT animal from allergic reaction.
PT
XX
XX Claim 8; Fig 11; 193pp; English.
XX
XX The invention provides a tertiary structure for the peanut allergen Ara H
CC 1. The Ara H 1 allergen is found to contain 23 linear IgE-binding
CC epitopes. The invention also provides an isolated recombinant peanut
CC allergen designated Ara h 3 and a nucleotide molecule encoding the peanut
CC allergen Ara h 3. Molecules of the invention are used to protect a host
CC animal from allergic reaction, particularly using a modified allergen
CC which is less reactive with IgE. The invention may also be used to ensure
CC that the allergen is not introduced into genetically modified food. The
CC present sequence represents a DNA encoding the Ara h 3 allergen. (Updated
CC on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 512 AA;
SQ
Alignment Scores:
Pred. No.: 3.64e-234 Length: 512
Score: 2652.00 Matches: 506
Percent Similarity: 99.4% Conservative: 0
Best Local Similarity: 99.4% Mismatches: 1
Query Match: 95.4% Indels: 2
DB: 2 Gaps: 1
US-10-728-323-3 (1-1524) x AAY40912 (1-512)
QY 1 CGGCAGCAACCGGAGGAGACGCTGCAGTTCAGCGCCTCAATGCGCAGACCTGAC 60
Dd 4 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 23
QY 61 AATCGCATTTGAATCAGAGGGCGGTTACATTGAGACTTTGGAAACCCCAACACAGGAGTTC 120
Dd 24 AsnArgIleGluSerGluGlyGlyTyrlleGluThrTrpAsnProAsnAsnGlnGluPhe 43

```

```

QY 121 GAATGCGCGCGTGCCTCTCTCGCTTAGTCTCGCGCGCAACGCCCTTCGTAGGCCT 180
Dd 44 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro 63
QY 181 TTCTACTCCAATGCTCCCGAGAGATCTTCATCCAGCAAGGAGGGGATACTTTGGTTG 240
Dd 64 PheTyrSerAsnAlaProGlnGluPheIleGlnGlnGlyArgGlyTyrPheGlyLeu 83
QY 241 ATATTCCCTGGTTGCTCTAGACACTATGAAGACCTCACACAAGGTGCTGATCTCAG 300
Dd 84 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 103
QY 301 TCCCAAGACCAACAGACGCTCTCCAAGGAGAGACCAAGCAACAGCAACAGATAGT 360
Dd 104 SerGlnArgProArgArgLeuGlnGlyGluAspGlnSerGlnGlnGlnArgAspSer 123
QY 361 CACCAGAAGGTGCACCGTTTCGATGAGGTGTCTCATTTGCAGTTCGCCCGGTGTGCT 420
Dd 124 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143
QY 421 TTCTGGCTCTACAACGACCAACGACACTGATGTTGTTGCTGTTCTTCTTACTGACCAAC 480
Dd 144 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 163
QY 481 AACACACGACACGACTTCATCAGTTCCCGAGAGATTCATTTGGCTGGACACGGAG 540
Dd 164 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 183
QY 541 CAA-----GAGTTCTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGAGCAAGAGCTTA 594
Dd 184 GlnGluPheGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgArgSerLeu 203
QY 595 CCATATAGCCCATACAGCCCGCAAGTCAAGCTAGACACAAGAGCGGTGAATTAGCCCT 654
Dd 204 ProTyrSerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerPro 223
QY 655 CGAGGACACACAGCGCGCAGAGAACGAGCAGGAGCAAGAGAAAGAAACCAAGGTGAAAC 714
Dd 224 ArgGlyGlnHisSerArgArgGluArgAlaGlyGlnGluGluGlnGluGlyGlyAsn 243
QY 715 ATCTTCAGCGGCTTCACGCGGAGTTCCTGGAAACAAGCCTTCCAGGTTGACGACAGACAG 774
Dd 244 IlePheSerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGln 263
QY 775 ATAGTCAAAAACCTATAGAGCGGAGACGAGAGTAGAAGAGAGCGGCATTTGACAGTG 834
Dd 264 IleValGlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrVal 283
QY 835 AGCGGAGCGCTCAGAACTTTAGCCCGCAGATAGAAAGAGACGTCGCCGACGAAGAGAGAA 894
Dd 284 ArgGlyGlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGlu 303
QY 895 TACGATGAAGATGAATATGAATACGATGAAGAGATAGAGGCGTGCAGGGGAAGCAGA 954
Dd 304 TyrAspGluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArg 323
QY 955 GGCAGGGGGAATCGTATTGAAGAGACGATCTGCACCGCAAGTCTTAAAGACATTTGGT 1014
Dd 324 GlyArgGlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGly 343
QY 1015 AGAAACAGATCCCTCAGCATCTACACCTCAAGCTGGTTTCACTCAAAACCTGCCAACGAT 1074
Dd 344 ArgAsnArgSerProAspIleTyrAsnProGlnAlaGlySerLeuIleThrAlaAsnAsp 363
QY 1075 CTCAACCTTCTAATACTTAGTGGCTTGGACCTAGTGTCTGAATATGAAATCTCTACAGG 1134
Dd 364 LeuAsnLeuLeuIleLeuArgTrpLeuGlyLeuSerAlaGluTyrGlyAsnLeuTyrArg 383
QY 1135 AATGCATTGTTTTCGCTCACTCAACACCAACGACACAGCATCATATATCGATTGAGG 1194
Dd 384 AsnAlaLeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArg 403

```


QY 1195 GGACGGCTCAGCTGCAAGTCGTGGACAGCAACCGCAACAGAGTGTACGACGAGGAGCTT 1254
 DB 404 GlyArgAlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeu 423
 QY 1255 CAAGAGGTCACGTGCTTGTGGTCCACAGAACTTCGCGTGCCTGGAAGTCCGAGAC 1314
 DB 424 GlnGluGlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSer 443
 QY 1315 GAGAACTTCGAATACCTGGGATTCAAGACAGACTCAAGGCCGAGCATAGCCAACTTCGCC 1374
 DB 444 GluAsnPheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAla 463
 QY 1375 GTGAAATCCGTCATAGATAACCTGCCGAGGAGTGGTTCGCAATTCATATGCGCTC 1434
 DB 464 GlyGluAsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeu 483
 QY 1435 CAAGGGACGACGACGAGCTTAAGAACACACACCCCTTCAAGTTCCTCGTTCACCG 1494
 DB 484 GlnArgGluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProPro 503
 QY 1495 TCTCAGCAGTCTCCGAGGCTGTGGCT 1521
 DB 504 SerGlnGlnSerProArgAlaValAla 512

RESULT 7

AAB33601

ID AAB33601 standard; protein; 526 AA.

XX AAB33601;

XX 12-SEP-2003 (revised)

DT 22-JAN-2001 (first entry)

XX Modified Ara h 3 amino acid sequence.

XX Allergy; allergic reaction; allergen; anaphylactic antigen; peanut;

KW Ara h 1; Ara h 2; Ara h 3; epitope; binding; immunoglobulin E; IgE;

KW food antigen; sensitising; immune response; anti-allergic.

XX Arachis hypogaea.

XX WO200051647-A2.

XX 08-SEP-2000.

XX 03-MAR-2000; 2000WO-US005655.

XX 03-MAR-1999; 99US-0122960P.

PR 06-DEC-1999; 99US-00455294.

XX (MOUN) MOUNT SINAI SCHOOL MEDICINE.

XX Sampson HA;

XX WPI; 2000-611341/58.

XX Non-human animal sensitized to an antigen, useful as an animal model for

PT studying allergic reactions to allergens, such as those in food and in

PT the environment.

XX Example 6; Fig 17C; 124pp; English.

XX The present invention describes an animal model which can be used for

XX studying allergic reactions to allergens. The animal is sensitized to a

CC selected antigen by administering the antigen itself or a nucleic acid

CC encoding the antigen, where preferably the antigen is an anaphylactic

CC antigen. The sensitized animal can then be used to screen for compounds

CC which may help to prevent, ameliorate, or cure allergic conditions in

CC humans. The animal model can be used for studying allergic reactions to

CC allergens, such as those in food (peanuts, fruit, berry, nut, bean, milk,

CC dairy products), or in the environment (weed pollen, grass pollen, tree

CC pollen, mite, animal, animal dander, fungal, and insect antigens).

CC AAB33478 to AAB33601 represent sequences which are used in examples from

CC the present invention to specifically examine the peanut allergy, and the
 CC peanut anaphylactic antigens Ara h 1 to 3. (Updated on 12-SEP-2003 to
 CC standardise OS field)

XX Sequence 526 AA;

Alignment Scores:
 Pred. No.: 7,12e-233 Length: 526
 Score: 2638.00 Matches: 502
 Percent Similarity: 99.0% Conservative: 0
 Best Local Similarity: 99.0% Mismatches: 5
 Query Match: 94.9% Indels: 0
 DB: 3 Gaps: 0

US-10-728-323-3 (1-1524) x AAB33601 (1-526)

QY 1 CGGCAGCAACCGGAGAGAACCGGTCCAGTTCAGCGCTCAATCCGCGAGACCTGAC 60
 DB 5 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 24
 QY 61 AATCCGATTGAATCAGAGGGCGTTACATTGAGACTTGGAAACCCCAACACAGGAGTTC 120
 DB 25 AsnArgIleGluSerGluGlyTyrIleGluThrTrpAsnAlaAsnAsnGlnGluPhe 44
 QY 121 GAATCGCGCGCTCGCCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCT 180
 DB 45 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro 64
 QY 181 TTCTACTCAATGCTCCCGAGAGATCTTCATCCAGCAGGAGGGGATACCTTGGTGTG 240
 DB 65 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 84
 QY 241 ATATTCCCTGTTGTCTAGACACTATGAAGAGCTTCACACAAAGTCTCGTATCTCAG 300
 DB 85 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgSerGln 104
 QY 301 TCCCAAGACCAACCAAGAGCTCTCAAGGAGAGAACCAAGCAACAGCAAGATAGT 360
 DB 105 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 124
 QY 361 CACGAGAGGTGCACCGTTCGATGAGGTGATCTCATTGCGATTCCTCCACCGGTGTGCT 420
 DB 125 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 144
 QY 421 TTCTGGCTCTACAACGACACGACACTGATGTTGTGTCTTCTTCTTACTCACACCAAC 480
 DB 145 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 164
 QY 481 AACAAACGACCAACCGCTTGCATCAGTTCCTCCAGGAGATTCAATTTGGCTGGGAACCGGAG 540
 DB 165 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 184
 QY 541 CAGAGTTCTTTAAGGTACCAACAAAGCAGACAAAGCAGACGAGAGCTTACCATAT 600
 DB 185 GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgSerLeuProTyr 204
 QY 601 AGCCCATACAGCCCGGAGTTCCTGGAAACAGCTTCCAGTTCAGCAGACAGATAGT 660
 DB 205 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 224
 QY 661 CAGCACGCCCGCAGAGAACGAGCAGCAAGAAAGAAAGAAACGAAAGTGGAAACATCTTC 720
 DB 225 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluAsnGluGlyGlyAsnIlePhe 244
 QY 721 AGCGGCTTCAGCCCGGAGTTCCTGGAAACAGCTTCCAGTTCAGCAGACAGATAGT 780
 DB 245 SerGlyPheThrProGluAlaLeuPheGlnAlaPheGlnValAspAspArgGlnIleVal 264
 QY 781 CAAACCTTAAGAGGGGAGACCGAGAGTGAAGAGAGGGAGCCATTTGTGACGTGAGGGGA 840
 DB 265 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 284
 QY 841 GGCCTCAGAATCTTGAGGCCACAGATAGAAAGAGACGTGCCGACGAAAGAGGAATACGAT 900

Db 285 GlyLeuArgAlaLeuSerProAspA:GlyArgAlaAspGluGluGluTyrAsp 304
QY 901 GAAGATCAATATCAATACGATGAAGAGATAGAGCGTGGCAGGGAAGCAGAGCAGG 960
Db 305 GluAspGluTyrAlaTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg 324
QY 961 GGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGCTTAAAGAACATTTGTTAGAAC 1020
Db 325 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaIleGlyAsnGlyArgAsn 344
QY 1021 AGATCCCTCAGATCTACACCTCAAGCTGCTCACTCAAACTGCCACGATCTCAAC 1080
Db 345 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 364
QY 1081 CTTCTAATCTAGTGGCTTGGACCTAGTCTGCAATATGGAATCTTACAGGAATGCA 1140
Db 365 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 384
QY 1141 TTGTTTGTCTCTACTACAAACACGACAGCATCATATATGATTGAGGGACGG 1200
Db 385 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 404
QY 1201 GCTCAGTCAAGTCTGGACAGCAAGCGCAACAGAGTGTCACGAGGAGCTTCAAGAG 1260
Db 405 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 424
QY 1261 GGTCACTGCTTGTGGTCCACAGAACTTCGCCGTGCTGGAAAGTCCACAGACGAGAAC 1320
Db 425 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 444
QY 1321 TTGGAATACGTGGATTCAAGACAGACTCAAGGCCAGCATAGCCAACTCGCGGTGAA 1380
Db 445 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 464
QY 1381 AACTCCGTATAGTAACTGCGGAGGAGTGTTCGCAATTCATATGCGCTCCAAAGG 1440
Db 465 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 484
QY 1441 GAGCAGCAAGGAGCTTAAGAACAAACACCCCTTCAAGTTCTTCGTTCCACCGTCTCAG 1500
Db 485 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProProSerGln 504
QY 1501 CAGTCTCCGAGGCTGTGGCT 1521
Db 505 GlnSerProArgAlaValAla 511

US-10-728-323-3 (1-1524) x AAU04711 (1-526)
QY 1 CGGCAGCAACCGGAGGAGAACGGTCCAGTTCCAGCGCTCAATGCGCAGACCTGAC 60
Db 5 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 24
QY 61 AATCGATTGAATCAGAGGGCGGTACATTGAGACTTGGAAACCCCAACACAGGAGTTC 120
Db 25 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnAlaAsnAsnGlnGluPhe 44
QY 121 GAATGCGCGGCTGCGCTCTCGCTAGTCTCCGCGCAACGCCCTTCGTAGGCGCT 180
Db 45 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro 64
QY 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGGATCTTTGGGTG 240
Db 65 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 84
QY 241 ATATTCCCTGGTTGCTCTAGACACTATGAAGACCTCACACAAAGTCTCGATCTCAG 300
Db 85 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgSerGln 104
QY 301 TCCCAAGACCCACCAAGACGCTCCAGGAGAGAGACCAAGCCAAACAGCAGATAGT 360
Db 105 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnGlnArgAspSer 124
QY 361 CACCAGAAGGTGACCGCTTTCGATGAGGGTGTCTCATTCAGTTCACCGGTGTGCT 420

PA (PANA-) PANACEA PHARM LLC.
PA (UYAR-) UNIV ARKANSAS.
XX (MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.
PI Bannon GA, Burks WA, Caplan MJ, Sampson H, Sosin H;
XX WPI; 2001-381378/40.
XX Antigenic fragments useful for reducing anaphylactic risk and reducing
PT the severity and/or number of allergic symptoms in individuals sensitive
PT to antigens, have reduced ability to bind Immunoglobulin E.
XX Disclosure; Fig 12; 100pp; English.
XX

CC The sequence represents the amino acid sequence of modified anaphylactic
CC antigen Ara h 3, which has been altered to disrupt immunoglobulin E (IgE)
CC binding sites. Ara h 3 is an anaphylactic antigen (A), which was used to
CC design antigenic peptides having a reduced ability to bind IgE as
CC compared with the intact (A), or having a sequence substantially
CC identical to a portion of sequence of an antigen that includes at least
CC one IgE binding site, where at least one IgE binding site of the peptide
CC is altered. The antigenic peptides are used in a composition which is
CC useful for reducing risk or severity of allergic reaction to an antigen.
CC This is done by identifying an individual at risk of allergic reaction to
CC an antigen by identifying prior display of allergic symptoms when exposed
CC to the antigen, or a familial relationship with an individual who
CC previously displayed allergic symptoms when exposed to the antigen.
CC Following this an antigen-specific IgE present on one or more mast cells
CC or basophils in the individual's serum is identified. The individual is
CC then contacted with a peptide corresponding to a portion of the antigen,
CC which is selected, formulated, and delivered so that binding of the
CC peptide to antigen-specific IgE is reduced as compared with IgE binding
CC of intact antigen. The composition is also useful for treating and
CC preventing allergic reactions
XX
SQ Sequence 526 AA;

Alignment Scores:
Pred. No.: 7.12e-233 Length: 526
Score: 2638.00 Matches: 502
Percent Similarity: 99.0% Conservative: 0
Best Local Similarity: 99.0% Mismatches: 5
Query Match: 94.9% Indels: 0
DB: 4 Gaps: 0

RESULT 8
AAU04711
ID AAU04711 standard; protein; 526 AA.
XX AAU04711;
XX AC
XX XX
DT 23-OCT-2001 (first entry)
XX
XX Modified anaphylactic antigen Ara h 3.
XX Ara h 3; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;
KW allergy; mast cell; basophil; mouse.
XX Mus sp.
OS Synthetic.
XX
PN WO200140264-A2.
XX
XX
PD 07-JUN-2001.
XX
XX 06-DEC-2000; 2000WO-US033124.
PF
XX 06-DEC-1999; 99US-00455294.
PR 23-JUN-2000; 2000US-0213765P.
PR 27-SEP-2000; 2000US-0235797P.
XX

Db 125 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 144
QY 421 TTCTGGCTCTACAAACACACGACACTGATGTTGTTGCTGTTCTTACTGACACCAAC 480
Db 145 PheTrpLeuTyrrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 164
QY 481 AACACGACAAACAGCTTTCAGTTCCCGAGAGATTCAATTTCGGCTCGGACACGGAG 540
Db 165 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 184
QY 541 CRAAGCTTCTTAGGTACCAAGCAACAAAGCAGACAGACAGAGAGAGCTTACCATAT 600
Db 185 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 204
QY 601 AGCCCATACAGCCCGCAAGTTCAGCTAGACAAAGAGAGCGTCAATTTAGCCCTCGAGA 660
Db 205 SerProTyr-SerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 224
QY 661 CAGCACAGCCGACAGAAACGAGCAGCAAGAAAGAAAGAAACGAAAGTGGAAACATCTTC 720
Db 225 GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIlePhe 244
QY 721 AGCGGCTTCACGCCGAGTTCCTGGAAACAAGCCTTCAGGTTGACGACAGACAGATG 780
Db 245 SerGlyPheThrProGluAlaLeuPheGlnAlaPheGlnValAspAspArgGlnIleVal 264
QY 781 CAAACCTAAGAGCCGAGACCGAGAGTGAAGAGAGGAGCCATTGTGACAGTGAAGGGA 840
Db 265 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 284
QY 841 GGCCTCAGAAATCTTGAGCCCGATAGAAAGAGACGTCGCCGACGAAAGAGGAATACGAT 900
Db 285 GlyLeuArgAlaLeuSerProAspArgLysArgAlaAspGluGluGluTyrAsp 304
QY 901 GAAGATCAATATCAATACATGAGAGAGATAGAGCGTGGCAGGGGAGAGAGAGCAGG 960
Db 305 GluAspGluTyrAlaTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg 324
QY 961 GGAATGGTATTGAAGAGACGATCTGCACCGAAGTGTAAAGAAACATTCGTGAAGAAC 1020
Db 325 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 344
QY 1021 AGATCCCTCGACATCTACAAACCTCAAGCTGTTCTCAAACTGCCCAACGATCTCAAC 1080
Db 345 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 364
QY 1081 CTTCTAATACTTAGTGGCTGGACCTAGTGTGTAATATGGAATCTTACAGGATGCA 1140
Db 365 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 384
QY 1141 TTGTTTGTGCTCACTACACCAACCAACGACGACATCATATATCGATTGAGGGGACGG 1200
Db 385 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 404
QY 1201 GTCACGTCGAAGTCTGGACAGCAACGGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
Db 405 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 424
QY 1261 GGTACGCTGTGTGTGTCACAGAACTTCGCGCTGCGTGGAAAGTCCAGAGCGAGAAC 1320
Db 425 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 444
QY 1321 TTCGAATAGTGCATTCAAGACAGACTCAAGCCCGACATAGCCAACTCCCGGTGAA 1380
Db 445 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 464
QY 1381 AACTCCGTCATAGATAACCTCGCGGAGAGGTGGTTGCAAAATTCATATGCGCTCCAAAGG 1440
Db 465 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 484
QY 1441 GAGCAGGAGGAGGAGTCTAAGAACAAACCCCTTCAGTCTTCTGTTCCACCGTCTGAG 1500
Db 485 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProProSerGln 504

QY 1501 CAGTCTCCGAGGCTGTGGCT 1521
Db 505 GlnSerProArgAlaValAla 511

RESULT 9

AAU05036
ID AAU05036 standard; protein; 526 AA.

AC AAU05036;

XX 24-OCT-2001 (first entry)

DT Modified anaphylactic peanut antigen Ara h 3.

DE Anaphylactic antigen; Ara h 3; peanut; desensitisation; antigen; allergy;
KW immunotherapy; rush immunotherapy; anaphylaxis; asthma; food allergen;
KW milk allergen; shellfish allergen; latex; drug; environmental allergen;
KW grass pollen; ovalbumin; insect venom; peanut oil; immunoglobulin E; IgE.

XX Arachis sp.

OS Synthetic.

XX WO200139799-A2.

PN 07-JUN-2001.

XX 06-DEC-2000; 2000WO-US033125.

XX 06-DEC-1999; 98US-00455294.

PR 23-JUN-2000; 2000US-0213765P.

PR 27-SEP-2000; 2000US-0235797P.

XX (PANA-) PANACEA PHARM LLC.

XX Caplan M;

PI WPI; 2001-408258/43.

DR Preventing allergic response against antigens, e.g. food and
PT environmental allergens such as peanut allergen or grass pollen, latex or
PT drug, comprises administering agent, e.g. a peptide, that blocks antigen
PT binding sites on offending IgE.

XX Disclosure; Fig 2; 76pp; English.

XX The sequence represents the amino acid sequence of modified anaphylactic
CC peanut antigen Ara h 3. The protein is used for active or passive
CC desensitisation of an individual to an antigen; for alleviating or
CC preventing allergic reactions and for decreasing the risk of allergic
CC reactions during immunotherapy or rush immunotherapy, anaphylaxis and
CC asthma. The antigen may be a food allergen (e.g. peanut or milk
CC allergen), shellfish allergen, environmental allergen (e.g. grass pollen
CC or tree pollen), latex, drug, pollen, ovalbumin, an insect venom antigen
CC or predominantly linear epitopes. The protein is useful for protecting an
CC individual against subsequent inadvertent or intentional exposure to
CC antigen, e.g. receiving blocking agent before eating a chocolate bar
CC which may inadvertently contain peanut components or before eating foods
CC prepared using peanut oil. Administration of the blocking agents does not
CC result in cross-linking of anti-antigenic immunoglobulin E (IgE). After
CC exposure to the agent, the individual's antigen sensitivity is at least
CC temporarily reduced. Only those IgE molecules that bind the offending
CC antigen and contribute to the risk of an allergic response are blocked

XX Sequence 526 AA;

Alignment Scores:

Pred. No.: 7,12e-233 Length: 526
Score: 2638.00 Matches: 502
Percent Similarity: 99.0% Conservative: 0
Best Local Similarity: 99.0% Mismatches: 5
Query Match: 94.9% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-3 (1-1524) x AAU05036 (1-526)		
QY	1	CGGCAGCAACCGGAGAGAACGGCTGCCAGTTCACAGCGCTCAATCGCAGAGACCTGAC
Db	5	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp
QY	61	AATCGATTGAATCAGAGGGCGTTTACATTGAGACTTGGAACCCCAACCAACGAGGATTC
Db	25	AsnArgIleGluSerGluGlyGlyTyriIeGluThrTrpAsnAlaAsnAsnGlnGluPhe
QY	121	GAATCGCGCGCTGCCCTCTCGCTTAGTCTCGCGCAACGCCCTTCCTAGGSCCT
Db	45	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro
QY	181	TTCTACTCCAAATGCTCCCAAGAGATCTTCATCCAGCAAGGAAGGGGATACTTTGGGTTG
Db	65	PheTy-SerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyriPheGlyLeu
QY	241	ATATTCCTCGTTGTCTAGACACTATGAAGACCTCAACACAAAGTGTCTCATCTCAG
Db	85	IlePheProGlyCysProArgHisTyriGluGluProHisThrGlnGlyArgArgSerGln
QY	301	TCCCAAGACCAACAGACGCTCCAGGAGAGACCAAGCCAAACAGCAACAGATAGT
Db	105	SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer
QY	361	CACCAAGGTGCACCGCTTTTCGATGAGGGTGATCTCATTCAGTTCCCAACCGGTGTTGCT
Db	125	HisGlnIlyValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla
QY	421	TTCTGGCTCTAACACGACCAACACTGATGTTGTTGCTGTTCTCTTACTGACCAAC
Db	145	PheTrpLeuTyriAsnAspHisAspThrAspValAlaValSerLeuThrAspThrAsn
QY	481	AACAAGCACAAACAGCTTGATAGTCCCGCAGAGATTCAATTTGGCTGGGAACACGGAG
Db	165	AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnThrGlu
QY	541	CAAGAGTCTTAAAGTACCAGCAACAAAGCAGACAAAGCAGACAGAAAGCTTACCATAT
Db	185	GlnGluPheLeuArgTyriGlnGlnSerArgGlnSerArgArgSerLeuProTyri
QY	601	AGCCATACAGCCGCGAAAGTCCAGCTAGACACAGAGAGCGTGAAATTAGCCCTCGAGGA
Db	205	SerProTyriSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly
QY	661	CAGCACAGCCGACAGACGAGGAGCAAGAAAGAAAGAAAGAAAGCAATCTTC
Db	225	GlnHisSerArgArgGluArgAlaGlyGlnGluGluGlnGluGlyAsnIlePhe
QY	721	AGCGGCTTCACGCGGAGTTCTCTGGMAACAAGCCCTTCCAGGTTGACGACAGACAGATGTG
Db	245	SerGlyPheThrProGluAlaLeuPheGlnAlaPheGlnValAspAspArgGlnIleVal
QY	781	CAAAACCTAAGAGCGAGACGAGAGTGAACAGAGGGGCCATTGTGACAGTGAGGGGA
Db	265	GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly
QY	841	GGCTCAGAAATCTTGAGCCAGATAGAAAGAGAGCGTCCCGACGAAGAGAGGAATACGAT
Db	285	GlyLeuArgAlaLeuSerProAspArgIlySargArgAlaAspGluGluGluTyriAsp
QY	901	GAAGATGAATATGAATACGATGAGAGGATAGAGCGGTGGCAGGGGAAGCAGAGCAGG
Db	305	GluAspGluTyriAlaTyriAspGluGluAspArgArgArgGlyArgGlySerArgGlyArg
QY	961	GGGAATCGTATTGAAGACAGATCTGCACCGCAAGTCTCTAAAGACATTTGGTAGAAC
Db	325	GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaIlySylsAsnIleGlyArgAsn
QY	1021	AGATCCCTCGACATCTCAACCCCTCAAGCTGGTTCTCACTCAAAACCTGCAACGATCTCAAC
US-10-728-323-3 (1-1524) x AAU05036 (1-526)		
Db	345	ArgSerProAspIleTyriAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn
QY	1081	CTTCTAAATACTTAGTGGCTTGAGACTAGTGTGTAATATGGAATCTCTACAGGAATGCA
Db	365	LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyriGlyAsnLeuTyriArgAsnAla
QY	1141	TTGTTTGTGCTCCTTACCAACACCAACGACACAGCATCATATATCGATTGAGGGACGG
Db	385	LeuPheValAlaHisTyriAsnThrAsnAlaHisSerIleIleTyriArgLeuArgGlyArg
QY	1201	GCTCAGCTCAAGTCTGTGACAGCAACCGCAACAGAGTGTACGACGAGAGCTTCAAGAG
Db	405	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyriAspGluGluLeuGlnGlu
QY	1261	GGTACGCTGCTGTGTCACAGAACTTCAGACAGACTCAAGGCCAGCATAGCCAACTCGCCGCTGAA
Db	425	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn
QY	1321	TTTCAATACGTGGCATTTCAAGACAGACTCAAGGCCAGCATAGCCAACTCGCCGCTGAA
Db	445	PheGluTyriValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu
QY	1381	AACTCCGTCATAGATAAATCTGCGGAGAGGTGGTTGCAAAATTCATATGCGCTCCAAAGG
Db	465	AsnSerValIleAspAsnLeuProGluGluValAlaAsnSerTyriGlyLeuGlnArg
QY	1441	GAGCAGCAAGGAGCTTAAAGCAACCAACCCCTTCAAGTCTTCTGTTCCACCGCTCTCAG
Db	485	GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPhePheValProProSerGln
QY	1501	CAGTCTCCGAGGCTGTGGCT
Db	505	GlnSerProArgAlaValAla
RESULT 10		
ID	ADO38357	
XX	ADO38357 standard; protein; 507 AA.	
AC	ADO38357;	
DT	15-JUL-2004 (first entry)	
XX	Peanut allergen Ara h 3.	
DE		
XX	Antibacterial; Virucide; Fungicide; Antiparasitic; Antiarthritic;	
KW	Antirheumatic; Neuroprotective; Antiinflammatory; Dermatological;	
KW	Immunosuppressive; Antidiabetic; Antithyroid; Antiasthmatic;	
KW	Antiallergic; Cytostatic; Antipsoriatic; Gene Therapy; Vaccine;	
KW	MHC Class II; Ii-key motif; immune response; anthrax; EBOLA; HIV;	
KW	influenza; vaccinia virus; infection; bacterium; virus; parasite; fungus;	
KW	rickettsia; rheumatoid arthritis; multiple sclerosis;	
KW	lupus erythematosus; diabetes mellitus; myasthenia gravis;	
KW	autoimmune thyroiditis; scleroderma; dermatomyositis; pemphigus; asthma;	
KW	allergic rhinitis; topical dermatitis; colitis; cancer; psoriasis;	
KW	adenoma; peanut; Ara h 3.	
XX		
OS	Arachis hypogaea.	
XX		
PN	US2004058881-A1.	
XX		
PD	25-MAR-2004.	
XX		
PF	24-SEP-2002; 2002US-00253286.	
XX		
PR	24-SEP-2002; 2002US-00253286.	
XX		
PA	(ANTI-) ANTIGEN EXPRESS INC.	
XX		
PI	Humphreys RE, Xu M;	
XX		
DR	WPI; 2004-294259/27.	
XX		
PT	New non-naturally occurring protein or polypeptide modified by	


```

|||||
43 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 62
|||||
121 GAATGCGCGGCTCGCTCTCTCGCTTAGTCTCCGCGCAACCGCCCTCGTAGGCT 180
|||||
63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 82
|||||
181 TTCTACTCCAATGCTCCCGAGAGATCTTCATCCAGCAAGGAGGGGATACTTTGGTTG 240
|||||
83 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 102
|||||
241 ATATTCCCTGGTGTCTAGACACTATGAAAGCCCTCACACAGGTCGTGATCTGAG 300
|||||
103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 119
|||||
301 TCCCAAGACCAACCAAGAGCTCTCCAAGGAGAAGACCAACCAAGCAGACGAGATGT 360
|||||
120 SerSerArgPro-----GlnAspArg 126
|||||
361 CACCAGAAGGTGCACCGTTTCGATGAGGTGATCTCATTTGCAGTTCCTCCACCGTGTGCT 420
|||||
127 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146
|||||
421 TTCTGCTCTACAAACGACACGACACTGATGTTGTTGCTGTTCTTCTTACTGACACCAAC 480
|||||
147 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 166
|||||
481 AACACGACGACACAGCTTCATGATTCCTCCAGGAGATTCAATTTGGCTGGGACACGGAG 540
|||||
167 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
|||||
541 CAAGAGTTCTTAAAGTACCAGCAACAAAGACGACAAAGACGAGCAGCAAGAGCTTACCATAT 600
|||||
187 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGly----- 199
|||||
601 AGCCCATACAGCCCGCAAGTCAAGCTAGACAAAGAGCGGTGAATTTAGCCCTCGAGGA 660
|||||
200 -----Gly 200
|||||
661 CAGCACAGCGCAGCAACGAGCAGGACGACGAAGAAGAAACGAAGCTGGAACATCTTC 720
|||||
201 ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluGluAsnGluGlyGlySerIleLeu 220
|||||
721 AGCGGCTTCACGCGGAGTTCTCTGGAACAAAGCCTTCAGGTTTGACACAGACAGATAGTG 780
|||||
221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 239
|||||
781 CAAACCTTAAGGCGCAGACCGAGACTGAAGAAGAGGGGCCATTTGTGACGTGAGGGA 840
|||||
240 ArgLysLeuGlnGlyGluAsnGluGluGluGluLysGlyAlaIleValThrValLysGly 259
|||||
841 GGCCTCAGAATCTTGAGCCCA-----GATAGAAAGAGACGTCGCCACGAAGAAGAG 891
|||||
260 GlyLeuSerValIleSerProThrGluGluGlnGlnGlnArgProGluGluGluGlu 279
|||||
892 GAATACGATCAAGATCAATATGAATACGATGAGAGGATAGAGGCGTGGCGAGGGAAC 951
|||||
280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294
|||||
952 AGAGCGAGGGGAATGGTATTGAAAGACAGATCTGCACCGCAAGTCTTAAAGAACATTT 1011
|||||
295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311
|||||
1012 GGTAGAAACAGATCCCTCAGATCTACACCTCAAGCTGGTTCTCACTCAAACTGCCAAC 1071
|||||
312 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrAlaThr 331
|||||
1072 GATCTCAACCTTCTAATCTTAGTGGCTGGACCTTAGTCTGCAATATGAAATCTCTAC 1131
|||||
332 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 351
|||||
1132 AGGAATGCATGTTGTTCTGCTCACTACAAACACCAACGACACGACGATCATATATCATGTTG 1191
|||||

```

```

352 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleTyrAlaLeu 371
1192 AGGGACCGGCTCAGCTGCAAGTCTGTCGACAGCAACGCGCAACAGAGTGTACACGAGGAG 1251
372 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 391
1252 TTTCAGAGGGTACGCTGCTTGTGTGTCACAGAACTTCGCGCTCGCTGGAAAGTCCCCAG 1311
392 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 411
1312 AGCAGAACTTCGATACGTCGTCATTAAGACAGACTCAAGCCCGACGATACCCACCTC 1371
412 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431
1372 GCGGTGAAAACCTCGTCATAGATAACCTGCGGAGAGGTTGGTTGCCAAATTCATATGCG 1431
432 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnThrPheAsn 451
1432 TTCCAAAGGAGCAGGCAAGCAGCTTAAGAAACAACAACCCCTTCAAGTTCCTGTTCCA 1491
452 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 471
1492 CCGTCTCAGCAGCTCTCCGAGGCGCTGTGGCT 1521
472 ProLysGluSerGlnArgValValAla 481

RESULT 14
ADG43988
ID ADG43988 standard; protein; 481 AA.
XX
AC ADG43988;
XX
DT 26-FEB-2004 (first entry)
XX
DE G. max glycinin G3 subunit protein.
XX
KW oil content; plant; storage protein; seed-specific promoter; 2S-albumin;
KW 7S-globulin; 11S-globulin; 12S-globulin; zein-prolamine; transgenic;
KW oil production; fat production; free fatty acid production; food;
KW animal feed; pharmaceutical; fine chemical production; glycinin.
XX
OS Glycine max.
XX
PN WO2003077643-A2.
XX
PD 25-SEP-2003.
XX
PF 17-MAR-2003; 2003WO-EP002733.
XX
PR 20-MAR-2002; 2002DE-01012893.
XX
PA (BADI ) BASF PLANT SCI GMBH.
XX
PI Bauer J;
XX
DR WPI; 2004-011485/01.
XX
N-PSDB; ADG43987.
PT Increasing total oil content of plants, useful e.g. as foods or animal
PT feeds, by reducing amount of storage proteins, particularly with double-
PT stranded interfering RNA.
PS Claim 4; SEQ ID NO 28; 253pp; German.
XX
CC This invention describes a novel method for increasing the total oil
CC content of a plant by reducing the amount of at least one storage protein
CC in the plant (or its tissue, organs, parts or cells) and selecting plants
CC that have higher total oil content than starting plants. The storage
CC protein is suppressed by introducing antisense RNA, optionally combined
CC with a ribozyme, sense RNA that induces co-suppression, DNA-binding
CC factors directed against storage protein genes, viral sequences that
CC degrade storage protein RNA, constructs that induce homologous
CC recombination of endogenous storage protein genes or mutations into

```

CC storage protein genes. Most preferably a plant cell is stably transfected
 CC with a recombinant expression construct, then regenerated to plants that
 CC express the incorporated sequence. The expression constructs particularly
 CC contain a seed-specific promoter and they are introduced into plants by
 CC standard methods, e.g. via Agrobacterium. The preferred storage proteins
 CC of the invention are 2S-albumens, 7S or 11S/12S-globulins or zein-
 CC prolamines. Transgenic organisms produced by the new method are used for
 CC production of oils, fats, free fatty acids or their derivatives, useful
 CC as foods, animal feeds, pharmaceuticals and fine chemicals. This sequence
 CC represents a storage protein used to illustrate the method of the
 CC invention.

XX Sequence 481 AA;

Alignment Scores:

Pred. No.: 1.5e-128 Length: 481
 Score: 1502.00 Matches: 297
 Percent Similarity: 71.6% Conservative: 68
 Best Local Similarity: 58.2% Mismatches: 91
 Query Match: 54.0% Indels: 54
 DB: 8 Gaps: 7

US-10-728-323-3 (1-1524) x ADG43988 (1-481)

QY 1 CGGCAGCAACCGGAGGAGACGGCTGCCAGTTCAGCGCTCAATGGCAGACCTGAC 60
 DB 23 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 42
 QY 61 AATCGCATTAATCAGAGCGCGTTCATTGACATTGGAACCCCAACACCAAGAGTTC 120
 DB 43 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 62
 QY 121 GAATGCCCGCGCTGCGCCTCTCTCGCTTAGTCTCGCGCGCAACGCCCTTCGTAGCCCT 180
 DB 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 82
 QY 181 TTCTACTCCCAATCTCCAGGAGATCTTCATCCAGCAGGAGGGGATACTTTGGTTG 240
 DB 83 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 102
 QY 241 ATATTCCTCGTTGTCTTAGACATATGAAGAGCCTCACACAAAGGTCTCGATCTCAG 300
 DB 103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 119
 QY 301 TCCCAAGACCAACAGACGCTCTCCAGGAGAGACCAAGCCCAACAGCAGATAGT 360
 DB 120 SerSerArgPro-----GlnAspArg 126
 QY 361 CACCAGAAGGTGCACCGTTTCGATGAGGGTGATCTCATTCAGTTCCTCCCGGTGTGCT 420
 DB 127 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146
 QY 421 TTCTGGCTTACCAACACCAACGACACTGATGTGTGTGCTGTTCTCTTACTGACCAAC 480
 DB 147 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 166
 QY 481 AACAAACGACCAACGAGTGTATAGTTCCTCCAGGAGATTCATTTGGCTGGGAACACGGAG 540
 DB 167 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
 QY 541 CAAGAGTTCCTTAAGGTACCAACAAAGCAGACAAAGCAGACGAGCAAGAGCTTACCATAT 600
 DB 187 GlnGluPheGluGlnTyrGlnProGlnLysGlnGlnGly----- 199
 QY 601 AGCCCCATACGCCCGCAAGTACGCTTAGCAAGAAGAGCGTGAATTTAGCCCTCGAGGA 660
 DB 200 -----Gly 200
 QY 661 CAGCAGCCCGCAGAGAACGAGCAGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
 DB 201 ThrGlnSerGlnLysGlyLysArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
 QY 721 AGCGGCTTCACGCGGAGTTCCTGGAAACAAGCCTTCCAGGTTGACGACGACAGATAGTG 780

DB 221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 239
 QY 781 CAAACCTTAAGAGCGGAGACCGAGAGTGAAGAGAGGAGGCCATTGTGACAGTAGGGGA 840
 DB 240 ArgLysLeuGlnGlyGluAsnGluGluGluGluLysGlyAlaIleValThrValLysGly 259
 QY 841 GGCCTCAGATCTTGAGCCCA-----GATAGAAAGAGACGTGCCGACGAAGAGAG 891
 DB 260 GlyLeuSerValIleSerProThrGluGluGlnGlnArgProGluGluGluGlu 279
 QY 892 GAATACGATCAAGATGAATATCAATACGATGAAGAGTAGAAGCGGTGCCAGGGCAAG 951
 DB 280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294
 QY 952 AGAGCAGGGGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTCTAAAAGAACATT 1011
 DB 295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311
 QY 1012 GGTAGAAACAGATCCCTGACATCTACAACTCAAGCTCAAGCTGCTCACTCAAACTGCCAAC 1071
 DB 312 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr 331
 QY 1072 GATCTCAACCTTCTAATCTTAGTGGCTTGGACCTTAGTGTCTCAATATGAAATCTCTAC 1131
 DB 332 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 351
 QY 1132 AGGAATGCATTTGTTGCTCGCTCACTACAAACACACGACACAGCATCATATATCGATTG 1191
 DB 352 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 371
 QY 1192 AGGGAGCGGCTCACGTGCAAGTCTGGACAGCAGCGCAACAGAGCTGTACGACGAGGAG 1251
 DB 372 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 391
 QY 1252 CTTCAAGAGCGTCACTGCTTGTGTGCCACAGAACTTCGCTCGCTCGGAAAGTCCCGAG 1311
 DB 392 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaArgSerGln 411
 QY 1312 AGCGAGAACTTCGAATACGTGCATTCAAGACAGACACTAAGGCCGACGATAGCAACCTC 1371
 DB 412 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431
 QY 1372 GCGGTGAACCTCCCTCATAGTAACCTGCCGAGGAGGTGTTGCTCAAAATTCATATGCC 1431
 DB 432 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 451
 QY 1432 CTCCAAAGGAGCAGGCAAGCGAGCTTAAGAAACAACCCCTTCAAGTTCTTCGTTCCA 1491
 DB 452 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 471
 QY 1492 CGCTCTCAGAGCTCTCCGAGGGCTGTGGCT 1521
 DB 472 ProLysGluSerGlnArgArgValValAla 481
 RESULT 15
 ADX74499
 ID ADX74499 standard; protein; 488 AA.
 XX
 AC ADX74499;
 XX
 DT 21-APR-2005 (first entry)
 XX
 DE Plant full length insert polypeptide seqid 43865.
 XX
 KW plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content.


```
||||| 399 LeuGlnGluGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 418
||||| 1312 AGCGAGAACTTCGAATACGTGGCATTCAAGACAGACTCAAGGCCAGCATAGCCAAACCTC 1371
||||| 419 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 438
||||| 1372 GCCGGTGAAACTCCGTCATAGATAACCTCCCGAGAGGTGGTTGCAAAATTCATATGGC 1431
||||| 439 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 458
||||| 1432 CTCCAAAGGAGCAGCAGGCGAGCTTAAGACAAACCCCTTCAAGTTCTTCGTTCCA 1491
||||| 459 LeuArgGlnGlnAlaArgGlnValLysAsnAsnProPheSerPheLeuValPro 478
||||| 1492 CCGTCTCAGCAGTCTCCGAGGGCTGTGGCT 1521
||||| 479 ProLysGluSerGlnArgArgValAla 488
```

Search completed: May 15, 2006, 21:45:05
Job time : 259.261 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 22:18:32 ; Search time 9.16611 Seconds
(without alignments)
4123.811 Million cell updates/sec

Title: US-10-728-323-3
Perfect score: 2779
Sequence: 1 cggcgagcaaccggaggagaa.....ctccgagggtgtggtcttaa 1524

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_n2p.model -DEV=xlp
-Q=/abss/ABSSWEB_spool/US10728323/runat_15052006_172137_22463/app_query.fasta_1
-DB=Issued Patents AA -OFFMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=Bits -START=1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs02p -USER=US10728323 @C@N_1_131 @runat_15052006_172137_22463
-NCPU=6 -ICPU=3 -NO MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/iaa/5 COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PTCUS COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/baCkfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1466	52.8	495	2	US-09-805-694B-16
2	811	29.2	492	2	US-09-462-720-2
3	623.5	22.4	483	2	US-10-053-410-6
4	293	10.5	141	2	US-09-645-593-12
5	227.5	8.2	85	2	US-09-645-593-10
6	215	7.7	165	2	US-09-645-593-11
7	184.5	6.6	489	2	US-09-424-283-3
8	173.5	6.2	96	2	US-09-645-593-9
9	172.5	6.2	626	2	US-09-106-872A-4
10	167	6.0	524	2	US-09-424-283-1
11	164	5.9	454	2	US-09-805-694B-4
12	160	5.8	448	2	US-09-323-195A-18

13	158	5.7	523	2	US-09-323-195A-17	Sequence 17, Appl
14	157.5	5.7	335	2	US-09-106-872A-17	Sequence 17, Appl
15	156.5	5.6	587	1	US-07-955-905A-23	Sequence 23, Appl
16	156	5.6	566	1	US-07-955-905A-2	Sequence 2, Appl
17	156	5.6	566	1	US-07-955-905A-22	Sequence 22, Appl
18	156	5.6	611	2	US-09-216-393B-81	Sequence 81, Appl
19	154.5	5.6	417	2	US-10-245-227D-14	Sequence 14, Appl
20	154.5	5.6	425	2	US-10-245-227D-12	Sequence 12, Appl
21	153.5	5.5	439	2	US-10-245-227D-1	Sequence 1, Appl
22	153.5	5.5	449	2	US-10-245-227D-89	Sequence 89, Appl
23	153.5	5.5	605	1	US-07-955-905A-24	Sequence 24, Appl
24	152	5.5	421	1	US-07-955-905A-27	Sequence 27, Appl
25	150	5.4	390	2	US-10-245-227D-2	Sequence 2, Appl
26	149.5	5.4	449	2	US-10-245-227D-95	Sequence 95, Appl
27	149.5	5.4	449	2	US-10-245-227D-97	Sequence 97, Appl
28	148.5	5.3	449	2	US-10-245-227D-91	Sequence 91, Appl
29	147.5	5.3	437	2	US-09-538-092-876	Sequence 876, App
30	147.5	5.3	614	4	PCT-US95-03236-21	Sequence 21, Appl
31	143.5	5.2	449	2	US-10-245-227D-93	Sequence 93, Appl
32	141.5	5.1	1898	1	US-08-056-200-94	Sequence 94, Appl
33	141.5	5.1	1898	1	US-08-800-644-94	Sequence 94, Appl
34	141.5	5.1	1898	2	US-09-538-092-1280	Sequence 1280, Ap
35	140.5	5.1	632	2	US-09-252-991A-25544	Sequence 25544, A
36	140	5.0	409	2	US-09-424-283-4	Sequence 4, Appl
37	139.5	5.0	481	2	US-09-252-991A-17707	Sequence 17707, A
38	139.5	5.0	1098	2	US-09-252-991A-20186	Sequence 20186, A
39	138	5.0	536	2	US-10-264-303-4	Sequence 4, Appl
40	137	4.9	540	2	US-10-264-303-3	Sequence 3, Appl
41	137	4.9	878	2	US-09-252-991A-31974	Sequence 31974, A
42	136	4.9	1250	1	US-08-441-139-9	Sequence 9, Appl
43	136	4.9	1250	2	US-09-487-558B-364	Sequence 364, App
44	135.5	4.9	1178	2	US-09-543-681A-7819	Sequence 7819, App
45	135	4.9	444	2	US-09-424-283-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-805-694B-16
; Sequence 16, Application US/09805694B
; Patent No. 6864362
; GENERAL INFORMATION:
; APPLICANT: Kinney, Anthony
; TITLE OF INVENTION: Hypoallergenic Transgenic Soybeans
; FILE REFERENCE: B01432 US NA
; CURRENT APPLICATION NUMBER: US/09/805,694B
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/189,823
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Glycine max
US-09-805-694B-16

Alignment Scores:	4.78e-133	Length:	495
Pred. No.:	1466.00	Matches:	288
Score:	70.1%	Conservative:	73
Percent Similarity:	55.9%	Mismatches:	104
Best Local Similarity:	52.8%	Indels:	50
Query Match:	2	Gaps:	6
DB:			
US-10-728-323-3 (1-1524) x US-09-805-694B-16 (1-495)			
Qy	1	CGCGAGCAACCGGAGAGAACCGGTGCCAGTTCACGCGCTCAATGCCGAGAGACTGAC	60
Db	23	ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProGly	42
Qy	61	AATCCATTGAATCAGAGCGGGTTCATTGAGATTGGAAACCCCAACACCGAGGATTC	120

```
Db 43 AsnArgIleGluSerGluGlyGlyLeuIleGluThrTrpAsnProAsnAsnLysProPhe 62
QY 121 GAATGCGCGCGTGCCTCTCTCGCTTAGTCTCTCGCGCGCAACGCCCTTCGTAGCCCT 180
Db 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 82
QY 181 TTCTACTCCAACTCTCCCGAGGAGATCTTCATCCAGCAGGAGGGGATACTTTGGTTG 240
Db 83 SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyLysGlyLeuPheGlyMet 102
QY 241 ATATTCCCTGGTGTCTCTAGACACTATGAAGAGCCTCACACAAAGCTCGTCATCTCAG 300
Db 103 IleTyrProGlyCysSerSerThrPheGluGluProGlnGlnArgGlyGln 122
QY 301 TCCCAAAGACCAACAAGCCTCTCCAAGGAGAGACCAAGCCAAACGACGAGATAGT 360
Db 123 SerSerArgPro-----GlnAspArg 129
QY 361 CACAGAGGTGCACCGCTTCGATGAGGTGATCTCATTTGCAGTTCCACCGGTGTGCT 420
Db 130 HisGlnLysIleTyrAsnSerArgGluGlyAspLeuIleAlaValProThrGlyValAla 149
QY 421 TTCTGGCTCTACAAACGACACGACACTGATGTTGTGCTGTTCTCTTACTACACCAAC 480
Db 150 TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerIleIleAspThrAsn 169
QY 481 AACCAACGACCAAGCTTGATCATGTTCCCGAGAGATTCATTTGGCTGGGAACACGGAG 540
Db 170 SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 189
QY 541 CAAGAGTCTTCAAGTACCAACGACCAACAAGCAGACCAAGCAGACGAGCAAGACCTTACCATAT 600
Db 190 GlnGluPheLeuLysTyrGlnGlnGln----- 199
QY 601 AGCCCATACAGCCGCAAGTCAGCTAGACCAAGAGCGTGAAATTTAGCCCTCGAGGA 660
Db 200 -----GlyGly 201
QY 661 CAGCACAGCGCAGAGAACGACGAGGACAGAGAACAGAAACGAAGTGGAAACATCTTC 720
Db 202 HisGlnSerGlnLysGlyLysHisGlnGlnGluGluAsnGluGlyGlySerIleLeu 221
QY 721 AGCGCTTACGCGGAGTCTCTGGACAAAGCTTCCAGCTTCAGCACACAGATAGTG 780
Db 222 SerGlyPheThrLeuGluPheLeuGluHisAlaPheSerVal---AspLysGlnIleAla 240
QY 781 CAAACCTTAAGAGCGCAGACCGAGAGTGAAGAAGAGGAGCGCATTTGTGACAGTGAAGGGA 840
Db 241 LysAsnLeuGlnGlyGluAsnGluGlyGluAspLysGlyAlaIleValThrValLysGly 260
QY 841 GGCTCTAGAATCTTGAGCCCA-----GATAGAAAGAGACGTGCGCAGCAGAAAGAG 891
Db 261 GlyLeuSerValIleLysProProThrAspGluGlnGlnArgProGlnGluGluGlu 280
QY 892 GAATACCATGAAGATGAATGAATAC-----GATGAAGAG---GATGAAGG 936
Db 281 GluGluGluGluAspGluLysProGlnCysLysGlyLysAspLysHisCysGlnArgPro 300
QY 937 CGTGGCAGGGAGACGAGGAGGAGGGAATGGTATTGAAGAGACGATCTGCACCGCAAGT 996
Db 301 ArgGlySerGlnSerLysSerArgArgAsnGlyIleAspGluThrIleCysThrMetArg 320
QY 997 GCTAAAGAACATTTGTAGAAACAGATCCCTGACATCTACAAACCTCAAGCTGGTTCA 1056
Db 321 LeuArgHisAsnIleGlyGlnThrSerSerProAspIleTyrAsnProGlnAlaGlySer 340
QY 1057 CTCAAAATGCGCAACGATCTCAACTCTTAATCTTAGTGGCTTGGACCTAGTGTGAA 1116
Db 341 ValThrThrAlaThrSerLeuAspPheProAlaLeuSerTrpLeuArgLeuSerAlaGly 360
QY 1117 TATGGAATCTCTACAGGAATGATGTTGTGCTCACTACACCAACGACACGACAGC 1176
Db 361 PheGlySerLeuArgLysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSer 380
```

```
QY 1177 ATCATATATCGATTGAGGGACGGGCTCAGTCCAGTCTGGACAGCAACGCAACAGA 1236
Db 381 IleIleTyrAlaLeuAsnGlyArgAlaLeuIleGlnValValAsnCysAsnGlyGluArg 400
QY 1237 GTGTACACAGAGAGCTTCAAGAGGTACAGTCTGTGTGTGTGTGTGTGTGTGTGTGT 1296
Db 401 ValPheAspGlyGluLeuGlnGluGlyArgValLeuIleValProGlnAsnPheValVal 420
QY 1297 GCTGGAAGTCCCGAGCAGAGCAATTCGAATACGTGGCATTTCAAGACAGACTCAAGGCC 1356
Db 421 AlaAlaArgSerGlnSerAspAsnPheGluTyrValSerPheLysThrAsnAspThrPro 440
QY 1357 AGCATAGCCAACTCCCGGTGAAACTCCGTATAGATAAATCTCCGGAGGAGGTGTT 1416
Db 441 MetIleGlyThrLeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIle 460
QY 1417 GCAAAATTCATATGCGCTCCAAAGGAGCAGGCAAGCAGCTTAAGAACCAACCCCTTC 1476
Db 461 GlnHisThrPheAsnLeuLysSerGlnGlnAlaArgGlnIleLysAsnAsnProPhe 480
QY 1477 AAGTTCTTCCATCCACCGTCTCAGCAGTCTCCGAGGGCTGTGGCT 1521
Db 481 LysPheLeuValProProGlnGluSerGlnLysArgAlaValAla 495
RESULT 2
US-09-462-720-2
; Sequence 2, Application US/09462720
; Patent No. 6617433
; GENERAL INFORMATION:
; APPLICANT: SOCIETE DES PRODUITS NESTLE
; APPLICANT: Marzaccini, Pierre
; APPLICANT: Rogers, John
; TITLE OF INVENTION: COFFEE STORAGE PROTEINS
; FILE REFERENCE: 8265-303
; CURRENT APPLICATION NUMBER: US/09/462,720
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: PCT/EP98/04038
; PRIOR FILING DATE: 1997-07-12
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Plant
US-09-462-720-2
Alignment Scores:
Pred. No.: 1.3e-69 Length: 492
Score: 811.00 Matches: 182
Percent Similarity: 52.4% Conservative: 89
Best Local Similarity: 35.2% Mismatches: 172
Query Match: 29.2% Indels: 74
DB: Gaps: 12
US-10-728-323-3 (1-1524) x US-09-462-720-2 (1-492)
```

```
QY 1 CGGACGCAACCGGAG-----GAGAACGCGTGCAGTTCCAGCGCCTCAATGCGCAG 51
Db 25 ArgProGlnProArgLeuArgGlyLysThrGlnCysAspIleGlnLysLeuAsnAlaGln 44
QY 52 AGACCTTGACAATCGCATTGAATCAGAGGGGGTTACATGAGACTTGGAAACCCCAACAAC 111
Db 45 GluProSerPheArgPheProSerGluAlaGlyLeuThrGluPheTrpAsnAsn 64
QY 112 CAGGAGTTCGAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTT 171
Db 65 ProGluPheGlyCysAlaGlyValGluPheGluArgAsnThrValGlnProLysGlyLeu 84
QY 172 CGTAGGCGCTTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAGGAGGGGATAC 231
Db 85 ArgLeuProHisTyrSerAsnValProLysPheValTyrValValGluGlyThrGlyVal 104
```

```
QY 232 TTGGTGTGATATTCCTCGTGTCTCTAGACACTATGAAGAGCCTCACACAAAGGTCTGT 291
Db   |||  ::  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
105 GlnGlyThrValIleProGlyCysAlaGluThrPheGlu-----SerGlnGlyGlu 121
QY 292 CGA-----TCTCAGTCCCAAGACACCAAGACGTCTCCAAGGAGAAGACCAA----- 339
Db   ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
122 SerPheTrpGlyGlnGluProGlyLysGlnGlyGlnGlnGlnGlnGlnGlnGlnGlnGln 141
QY 340 -----AGCCAAACAGCAACGA-----GATAGTCACCAGAAAGGTGCACCGT 378
Db   ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
142 LysGlyGlyGlnGluGlnArgGlnArgPheProAspArgHisGlnLysLeuArgArg 161
QY 379 TTCGATGAGGGTATCTCATGTCAGTCCACCGGTGTTCTTCTGCTCTACACAGAC 438
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
162 PheGlnLysGlyAspValLeuLeuLeuLeuProGlyPheThrGlnTrpThrTrpAsnAsp 181
QY 439 CAGCAGCTGATGTTGTTGTTCTCTTCTTCTACTGACACCAACAAACAGCAACAGCTT 498
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
182 GlyAspValProLeuValThrValAlaLeuLeuAspValAlaAsnGluAlaAsnGlnLeu 201
QY 499 GATCAGTTCCTCCAGGAGATTCAATTTGGCTGGGAACACGAGCAGAGTTCCTTAAGGTAC 558
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
202 AspLeuGlnSerLysLysPheLeuAlaGlyAsnProGlnGlnGlnGlyLysGlu 221
QY 559 CAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATATAGCCCATACGCCCGCAA 618
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
222 GlyHisGlnGlnGlnGlnHisArg----- 231
QY 619 AGTCAGCCTAGACNAGAAGCGTGAATTTAGCCCTCGAGGACAGCAGCCGCGAGNA 678
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
231 ----- 231
QY 679 CGAGCAGCACAAGAAGAAAGAAAGGTTGGAACATCTTCAGCGGCTTCACGCCGGAG 738
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
232 -----AsnIlePheSerGlyPheAspGln 240
QY 739 TTCCTGGAACAAGCTTCAGGTTGACGACAGACAGATAGTGCAAAACCTAAGAGCGGAG 798
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
241 LeuLeuAlaAspAlaPheAsnVal---AspLeuLysIleIleGlnLysLeuLysGly--- 258
QY 799 ACCGAGAGTGAAGNAGAGGAGCCATTGTGCAGTGTAGGGGA---GGCCTCAGATCTTG 855
Db   ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
259 ---ProLysAspGlnArgGlySer-----ThrValArgAlaGluLysLeuGlnLeuPhe 275
QY 856 AGCCACAGATAGAAGAGAGCTGCCGACGAAAGAGAGGAATACGATCAAGATCAATGAA 915
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
276 LeuProGlyTrpSerGluGlnValGlnGlnProGlnGln----- 288
QY 916 TAGCATGAAGAGATAGAAGCGGTGGCAGGGGAAGCAGAGCGGGGAATGTTATTGAA 975
Db   ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
289 ---GlnGlnGluGlnGlnGlnHisGlyValGlyValGlyArgGlyTrpArgSerAsnGlyLeuGlu 307
QY 976 GAGACGATCTGCACCGCAAGTGTAAAGAGACATTTGTTAGAAACAGATCCCTCATCATC 1035
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
308 GluThrLeuCysThrValLysLeuSerGluAsnIleGlyLeuProGlnGlnAlaAspVal 327
QY 1036 TACAACCTCAAGCTGGTTCATCAAACTGCAACAGCATCTCAACCTCTTAATCTTAGG 1095
Db   ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
328 PheAsnProArgAlaGlyArgIleThrThrValAsnSerGlnLysIleProIleLeuSer 347
QY 1096 TGGCTTGGACCTAGTGTGAATATGGAATCTCTACAGGAATGCATTTGTTGCTCATC 1155
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
348 SerLeuGlnLeuSerAlaGluArgGlyPheLeuTrpSerAsnAlaIlePheAlaProHis 367
QY 1156 TACAACACCAACGCACACAGCATCATATATTCATTGAGGGGACGGCTCACGTGCAAGTC 1215
Db   ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
368 TrpAsnIleAsnAlaHisAsnAlaLeuTrpValIleArgGlyAsnAlaArgIleGlnVal 387
QY 1216 GTGGCAGCAGCAACGAGAGTGTACGAGGAGGACTTCAAGAGGTCACGTGTTGTG 1275
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
388 ValAspHisLysGlyAsnLysValPheAspAspGluValLysGlnGlyGlnLeuIle 407
QY 1276 GTGCCACAGAACTTCGCCCGTCTGCGAAAGTCCCAGAGCGAGAACTTCGAATACGTGGCA 1335
```

```
Db 408 ValProGlnTrpPheAlaValIleLysLysAlaGlyAsnGlnGlnGlyPheGluTrpValAla 427
QY 1336 TTCACAGACAGACTCAAGGCCCCAGCATACCCACCTCGCGGTGAAACTCCGTCTAGAT 1395
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
428 PheLysThrAsnAspAsnAlaMetIleAsnProLeuValGlyArgLeuSerAlaPheArg 447
QY 1396 AACCTGCGGAGGAGGTGTTGCCAAATTCATATGCGCTCCAAAGGGAGCAGGCAAGGCAG 1455
Db   ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
448 AlaIleProGluGluValLeuArgSerPheGlnIleSerSerGluGluAlaGluGlu 467
QY 1456 CTTAAGAACAACAACCCCTTCAAGTTCCTGTTCCACCGTCTCAGCAGTCT 1506
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
468 LeuLysTrpGlyArgGlnGluArgLeuLeu-----SerGluGlnSer 482

RESULT 3
US-10-053-410-6
; Sequence 6, Application US/10053410
; Patent No. 6858778
; GENERAL INFORMATION:
; APPLICANT: Sewalt, Vincent J. H.
; APPLICANT: Jung, Rudolf
; APPLICANT: Hu, Wang-Nan
; APPLICANT: Meeley, Robert B.
; TITLE OF INVENTION: Improved Grain Quality Through Altered
; FILE REFERENCE: Expression of Seed Proteins
; CURRENT APPLICATION NUMBER: US/10/053,410
; CURRENT FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,455
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Zea mays
US-10-053-410-6

Alignment Scores:
Pred. No.: 1,86e-51 Length: 483
Score: 623.50 Matches: 151
Percent Similarity: 45.1% Conservative: 77
Best Local Similarity: 29.9% Mismatches: 172
Query Match: 22.4% Indels: 105
DB: 2 Gaps: 11

US-10-728-323-3 (1-1524) x US-10-053-410-6 (1-483)
QY 25 TGCCAGTTTCCAG---CGCCTCAATGCGCAGAGACCTGACAATCGCATTTGAATCAGAGGC 81
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
48 CysGlyPheAspGlyLysLeuGluAlaLeuGluProArgHisLysValGlnSerGluAla 67
QY 82 GGT-----TACATTGAGACTTGGAAACCCCAACACACAGGAGTTTCAATGCCCGGC 132
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
68 GlySerValGlnTrpPheSerArgPheAsnGluAlaAspArgGluLeuThrCysAlaGly 87
QY 133 GTCGCGCTCTCTCGCTTAGTCTCTCGCGCAACCGCCTTCGTAGSCCTTTCTACTCCAAT 192
Db   ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
88 IlePheAlaValArgValValValAspAlaMetGlyLeuLeuLeuProArgTrpSerAsn 107
QY 193 GCTCCCAAGGAGATCTTCATCCAGCAAGGAGGGGATACCTTTGGTGTGATTTCCCTCGT 252
Db   ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
108 ValHisSerLeuValTyrlleValGlnGlyArgGlyIlelleGlyPheSerPheProGly 127
QY 253 TGCTCTAGACACTATGAAGAGCCTCACACAAAGTCTCGATCTCAGTCCCAAGACCA 312
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
128 Cys----- 128
QY 313 CCAAGACGTCTCCAAGAGAGAACCAAGCCCAACAG----- 348
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
129 -----GlnGluGluThrGlnGlnGlnTrpGlyTrpGlyTyrGlyTyrGly 144
```


; APPLICANT: Moloney, Maurice
; APPLICANT: Singh, Surinder
; TITLE OF INVENTION: Flax Seed Specific Promoters
; FILE REFERENCE: 9369-151
; CURRENT APPLICATION NUMBER: US/09/645,593
; CURRENT FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/151,044
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: US 60/161,722
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Linum usitatissimum
; SEQ ID NO 10

Alignment Scores:
Pred. No.: 1-69e-13 Length: 85
Score: 227.50 Matches: 46
Percent Similarity: 61.2% Conservative: 14
Best Local Similarity: 46.9% Mismatches: 25
Query Match: 8.2% Indels: 13
DB: 2 Gaps: 2

US-10-728-323-3 (1-1524) x US-09-645-593-10 (1-85)

QY 220 GGAAGGGATCTTTGGTGTGATATTCCTGGTGTCTAGACACTATGAAGAGCTCTAC 279
Db 1 GlyArgGlyValThrGlyIleMetPhePro***CysProGluThrPheGluGluSerGln 20
QY 280 ACACAGGTCGTCGATCTCAGTCCCAAGACACCAAGACGCTCCAGGAGAACCA 339
Db 21 GlnGlnGlyGlnGlnGln-----GlnGlySerSerGln 32
QY 340 AGCCAAACAGACGAGATAGTCACCAAGAGTGCACCGTTTCGATGAGGTCATCTCATT 399
Db 33 -----AspGlnHisGlnIlyIleArgGlyPheArgGlyValIle 47
QY 400 GAGTTCACCGGTCCTTCTGCTCTACAGCACACGACACTGATGTTGCT 459
Db 48 AlaValProAlaGlyValAlaHisTrpSerTyrAsnAspGlyAsnGluProValMetAla 67
QY 460 GTTCTCTTACTCACACCAACACACGACACCGCTTGATCAGTTCCTCCAGG 513
Db 68 IleValValHisAspThrSerSerHisLeuAsnGlnLeuAspAsnProArg 85

RESULT 6
US-09-645-593-11
; Sequence 11, Application US/09645593
; Patent No. 6777591
; GENERAL INFORMATION:
; APPLICANT: Chaudhary, Sarita
; APPLICANT: van Rooijen, Gijb
; APPLICANT: Moloney, Maurice
; APPLICANT: Singh, Surinder
; TITLE OF INVENTION: Flax Seed Specific Promoters
; FILE REFERENCE: 9369-151
; CURRENT APPLICATION NUMBER: US/09/645,593

; CURRENT FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/151,044
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: US 60/161,722
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Linum usitatissimum
; SEQ-09-645-593-11

Alignment Scores:
Pred. No.: 3-89e-12 Length: 165
Score: 215.00 Matches: 62
Percent Similarity: 43.1% Conservative: 28
Best Local Similarity: 29.7% Mismatches: 69
Query Match: 7.7% Indels: 50
DB: 2 Gaps: 6

US-10-728-323-3 (1-1524) x US-09-645-593-11 (1-165)

QY 517 TTCAATTTGGCTGGGAACACGAGCAAGAGTTCTTAAGGTACCAACAAAGCAGACAA 576
Db 2 PheTyrLeuAlaGlyAsnProArgAspGluPheGluGlnSerGlnGlnGlyGlyArgLeu 21
QY 577 AGCAGACCAAGAGCTTACCATATAGCCCATACAGCCCGCAAGGTCAGCCTAGACA 636
Db 22 SerArgGlyGluSer----- 26
QY 637 GAGCGTGAATTTAGCCCTCGAGGACAGCACAGCCGACAGAGAGGAGGAGCAAGA 696
Db 27 -----GluGlyGlyArgGlyArgGluProLeuGlnProAlaThr 40
QY 697 GAAACAGAGGTGGAACATCTTCAGCGCTTCACCGGAGTTCTCGAACAAAGCCTTC 756
Db 41 ThrSerSer-----CysGlyIleAspSerLysLeuIleAlaGluAlaPhe 55
QY 757 CAGGTTGACGACAGACAGATAGTCAAAACCTAGAGCGGACGACGAGTGAAGAGAG 816
Db 56 AsnValAspGlu---AsnValAlaArgArgLeu-----GlnSerGluAsnAspAsnArg 72
QY 817 GGAGCCATTGTGACAGTGAAGGGGGCTCAGAATCTTGAGCCGACAGATAGAAAGAGAC 876
Db 73 GlyGlnIleValArgValGluGlyGluLeuAspIleValArgProThrSerIleGln 92
QY 877 GCCGACGAAGAGGAATACGATGAAGATGAATATGAATACGATGAAGAGGATGAAGG 936
Db 93 GluGluSerGlnGluGln----- 98
QY 937 CGTGCAGGGGACAGAGGAGGAGGGGG-----AATGGTATTGAGAGAGAGATC 984
Db 99 -----GlyGlyArgGlyGlyGlyArgTyrTyrSerAsnGlyValGluGluThrPhe 115
QY 985 TGCACCGCAAGTGTCTAAAGAAACATTCGTAGAAACAGATCCCTCGACATCTACACCT 1044
Db 116 CysSerMetArgLeuIleGluAsnIleGlyAspProSerArgAlaAspIlePheThrPro 135
QY 1045 CAAGCTGTTCTACCAAACTGCCAACGATCTCAACCTTCTTAATCTTAGTGGCTTGA 1104
Db 136 GluAlaGlyArgValArgSerLeuAsnSerHisLeuProValLeuGlnTrpIleGln 155
QY 1105 CTTAGTCTGAATATGGAATCTCTAC 1131
Db 156 LeuSerAlaGluArgGlyValLeuTyr 164

RESULT 7
US-09-424-283-3
; Sequence 3, Application US/09424283
; Patent No. 6437219
; GENERAL INFORMATION:
; APPLICANT: Grimes, et al.

RESULT 8
US-09-645-593-9
; Sequence 9, Application US/09645593
; Patent No. 6777591
; GENERAL INFORMATION:
; APPLICANT: Chaudhary, Sarita
; APPLICANT: van Rooijen, Gijs
; APPLICANT: Moloney, Maurice

APPLICANT: Singh, Surinder
TITLE OF INVENTION: Flax Seed Specific Promoters
FILE REFERENCE: 9369-151
CURRENT APPLICATION NUMBER: US/09/645,593
PRIORITY FILING DATE: 2000-08-25
PRIOR APPLICATION NUMBER: US 60/151,044
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: US 60/161,722
PRIOR FILING DATE: 1999-10-27
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 96
TYPE: PRT
ORGANISM: Linum usitatissimum
US-09-645-593-9

Alignment Scores:
Pred. No.: 3,06e-08 Length: 96
Score: 173.50 Matches: 35
Percent Similarity: 63.5% Conservative: 12
Best Local Similarity: 47.3% Mismatches: 24
Query Match: 6.2% Indels: 3
DB: 2 Gaps: 2

US-10-728-323-3 (1-1524) x US-09-645-593-9 (1-96)

QY 1 CGGACGACCGGAGGAG--AACGGTCCAGCTTCCAGCGCTCAATCGCGAGACCT 57
DB 25 ArgGlnPheGlnGlnGlnGlnCysGlnIleAspArgIleAspIleAspGluPro 44
QY 58 GCAATCGCATTAATCAGAGCGCGGTTACATTGAGACTTGAACCCCAACACGAGGAG 117
DB 45 AspLysThrIleGlnAlaGluAlaGly-----GluValTrpAspGlnAsnArgGlnGln 62
QY 118 TTCGATGCGCGCGGCTCCCTCTCTCGTTAGTCTCCGCGCGCAACGCCCTTCGTAGG 177
DB 63 PheGlnCysAlaGlyValAlaValArgThrIleGluProLysGlyLeuLeuLeu 82
QY 178 CTTTCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAA 219
DB 83 PropheTySerAsnThrProGlnLeuIleTyIleValGln 96

RESULT 9

US-09-106-872A-4
Sequence 4, Application US/09106872A
Patent No. 6486311
GENERAL INFORMATION:
APPLICANT: Burks Jr., A. Wesley
APPLICANT: Stanley, J. Steven
APPLICANT: Cockrell, Gael
APPLICANT: King, Nina E.
APPLICANT: Sampson, Hugh A.
APPLICANT: Helm, Ricki M.
APPLICANT: Bannon, Gary A.
TITLE OF INVENTION: Peanut Allergens and Methods
FILE REFERENCE: HS 103 CIP
CURRENT APPLICATION NUMBER: US/09/106,872A
CURRENT FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: PCR/US96/15222
PRIOR FILING DATE: 1996-09-23
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 626
TYPE: PRT
ORGANISM: Arachis hypogaea
FEATURE:
OTHER INFORMATION: Amino Acids 25-34 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 1
OTHER INFORMATION: Amino Acids 48-57 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 2
OTHER INFORMATION: Amino Acids 65-74 are Ara H 1 binding epitope,

OTHER INFORMATION: peptide 3
OTHER INFORMATION: Amino Acids 89-98 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 4
OTHER INFORMATION: Amino Acids 97-106 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 5
OTHER INFORMATION: Amino Acids 107-116 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 6
OTHER INFORMATION: Amino Acids 123-132 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 7
OTHER INFORMATION: Amino Acids 134-143 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 8
OTHER INFORMATION: Amino Acids 143-152 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 9
OTHER INFORMATION: Amino Acids 294-303 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 10
OTHER INFORMATION: Amino Acids 311-320 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 11
OTHER INFORMATION: Amino Acids 325-334 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 12
OTHER INFORMATION: Amino Acids 344-353 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 13
OTHER INFORMATION: Amino Acids 393-402 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 14
OTHER INFORMATION: Amino Acids 409-418 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 15
OTHER INFORMATION: Amino Acids 461-470 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 16
OTHER INFORMATION: Amino Acids 498-507 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 17
OTHER INFORMATION: Amino Acids 525-534 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 18
OTHER INFORMATION: Amino acids 539-548 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 19
OTHER INFORMATION: Amino acids 551-560 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 20
OTHER INFORMATION: Amino acids 559-568 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 21
OTHER INFORMATION: Amino acids 578-587 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 22
OTHER INFORMATION: Amino acids 597-606 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 23
US-09-106-872A-4

Alignment Scores:

Pred. No.: 1.02e-07 Length: 626
Score: 172.50 Matches: 117
Percent Similarity: 35.2% Conservative: 82
Best Local Similarity: 20.7% Mismatches: 204
Query Match: 6.2% Indels: 163
DB: 2 Gaps: 25

US-10-728-323-3 (1-1524) x US-09-106-872A-4 (1-626)

QY 4 CAGCAACCGAGGAGAACGGTCCAGCTTCCAGCGCTCAATGCGCAGAGACCTGCAAT 63
DB 130 ArgGlnProArgLysP-----TrpArgProSerHisGlnGlnPro---Arg 145
QY 64 CGCATTGAATCAGAGCGCGGTTACATTGAGACTTGAACCCCAACACGAGGATTCGAA 123
DB 146 LysIleArgProGluGly-----ArgGluGlyGluGlnGlnIleTrpGly 159
QY 124 TCGCGCGCGTCCGCTCTCTCGTTAGTCTTCCGCGCAACGCCCTTCGTAGGCCCTTTC 183
DB 160 ThrProGlySerHisValArgGluGluThrSerArgAsnAsn-Prophe----- 175
QY 184 TACTCCAATGTCCCGCAGGAGATCTTCATCCAGCAAGGAGGGATACTTGGTGTGATA 243
DB 176 -----TyrPheProSerArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgI 193
QY 244 TTCCTCGTGTGCTCTAGACACTATCAAGAGC-----CTCACACACAA 285
DB 193 eArgValLeuGlnArgPheAspGlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisar 213

```
QY 286 GGTCTCGATCTCAGTCCCAAA-----GACCACCAAGACGTCTC 324
Db : : : : :
QY 213 gileValGlnIleGluAlaLysProAsnThrLeuValLeuProLysHisAlaAspAlaAs 233
Db : : : : :
QY 325 -----CAAGGAGAAGACCAAAAGCCCAACAGCAACGAGATAGTCAACCA 365
Db : : : : :
QY 233 pAsnIleLeuValIleGlnGlnGlyGlnAlaThrValThrValAlaAsnGlyAsnAsnAr 253
Db : : : : :
QY 366 GAAGGTGCACCGTTTCGATGAGGTGATCTCAATTGCAGTTCGCCACCGGTGTTGCTTCG 425
Db : : : : :
QY 253 glysSerPheAsnLeuAspGluGlyHisAlaLeuArgIleProSerGlyPheIleSerTy 273
Db : : : : :
QY 426 GCTCTACACGACCAACGACACTGATGTTGTTGCTGTTCTCT---CTTACTGACACCAACAA 482
Db : : : : :
QY 273 rIleLeuAsnArgHisAsnAsnGlnAsnLeuArgValAlaLysIleSerMetProValAs 293
Db : : : : :
QY 483 CAACGCAACACCGATGATCAGTTCCTCCAGGAGATTCATTTGGCTGGGAACACGAGCA 542
Db : : : : :
QY 293 nThrProGlyGlnPheGluAspPhe-----PheProAlaSerSerArgAspGlnSe 310
Db : : : : :
QY 543 AGAGTTCCTTAAGGTACAGCAACAAAGACAGCAAAAGCAGACGAAAGAGCTTACCATATAG 602
Db : : : : :
QY 310 rSerTyrLeu-----GlnGluPheSerArgAsnThrLeu----- 321
Db : : : : :
QY 603 CCCATACACGCCCAAGTCAGCTAGACAAAGAGCGTGAATTTAGCCCTCGAGGACA 562
Db : : : : :
QY 663 GCACAGCCGCAGA-----CAACGAGCAGCAACAAAGAAAGAAACCAAGAGTGG 710
Db : : : : :
QY 330 nGluIleArgArgValLeuLeuGluAsnAlaGlyGlyGlnGlnGluArgGly-- 349
Db : : : : :
QY 711 AAACATCTTCAGCGGCTTCACGCCGGAGTTCCTGGAACAAGCCTCCAGGTTGACACAG 770
Db : : : : :
QY 349 ----- 349
Db : : : : :
QY 771 ACNAGATAGTCAAAACCTAAGAGCGCAGACCGAGAGTGAAGAA-----GAGGAGCCAT 824
Db : : : : :
QY 350 -----GlnArgArgTrpSerThrArgSerSerGluAsnAsnGluGlyValI 365
Db : : : : :
QY 825 TGTGACAGTGAGGGGAGGCTCAGAATCTTGAGCCCAAGATAGAAAGAGAGAGCTGCCACGA 884
Db : : : : :
QY 365 eValLysVal----- 368
Db : : : : :
QY 885 AGAAGAGGAATACGATGAGATGAATATGAATACGATGAAGAGATAGAAAGCGTGGCAG 944
Db : : : : :
QY 369 -SerLysGluHisValGluGluLeuThrLysHisAlaLysSerValSerLysLysGlySe 388
Db : : : : :
QY 945 GCGAAGCAGAGCGAGGGGGAATGGTATTGAAGAGCAGATCTGCACCCCAAGTCTAAAAA 1004
Db : : : : :
QY 388 rGluGluGluGly-----AspIleThrAsnProIleAsnLeuArgGluGlyGluPr 405
Db : : : : :
QY 1005 GAACATTGGTAGAAAC-----AGATCCCTCGACATCTACAAACCTCA 1046
Db : : : : :
QY 405 oAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAspLysLysAsnProGl 425
Db : : : : :
QY 1047 AGCTGGTTCACTCAAAACGTCACCAACGATCTCAACCTCTTAATACTTAGTGGCTTGGACC 1106
Db : : : : :
QY 425 n-----LeuGlnAspLeuAspMetMetLeu----- 433
Db : : : : :
QY 1107 TAGTGTGTAATATGGAATCTTACAGGAATGATGTTGTCGCTCACTACAAACACCAA 1166
Db : : : : :
QY 434 -----ThrCysValGluIleLysGluGlyAlaLeuMetLeuProHisPheAsnSerly 451
Db : : : : :
QY 1167 CGCACACAGATCATATATCGATTGAGGGGAGCGGCTCAGTGCAAGTCTGCTG----- 1218
Db : : : : :
QY 451 sAlaMetValIleValValValAsnLysGlyThrGlyAsnLeuGluLeuValAlaValAr 471
Db : : : : :
QY 1219 -----GACAGCAAA 1226
Db : : : : :
QY 471 glysGluGlnGlnArgGlyArgArgGluGluGluGluAspGluGluGlu 491
Db : : : : :
QY 1227 CGGCAAC-----AGATGTACGACGAGGAGCTTCAAGAGGGTCAACGTGCTGT 1274
```

```
Db 491 uGlySerAsnArgGluValArgArgTyrThrAlaargLeuLysGluGlyAspValPheI 511
QY 1275 GGTCCACACAGAACTTTCGCGTCGTCGGAAGTCCAGAGCGAGAACTTCGAATACGTGC 1334
Db : : : : :
QY 511 eMetProAlaAlaHisProValAlaIleAsnAlaSerSerGlu---LeuHisLeuLeuGl 530
Db : : : : :
QY 1335 ATTC-----AAGACAGACTCAAGGCCCCAGCATAGCAACCTCGCGGTGAAACTCCGT 1388
Db : : : : :
QY 530 yPheGlyIleAsnAlaGluAsnAsnHisArgIlePheLeuAlaGlyAspLysAspAsnVa 550
Db : : : : :
QY 1389 CATAGATAACCTGCGGAGGAGTGGTTCGAAATTCATATTCATATTCGCTCCAAAGGAGGACGC 1448
Db : : : : :
QY 550 lIleAspGlnIleGluLysGlnAlaLysAspLeuAlaPheProGlySerGlyGluGlnVa 570
Db : : : : :
QY 1449 AAGCGAGCTTAAGAAACAACACCCCTTCAAGTTCCTTCGTT-----CCACCGTCTCA 1499
Db : : : : :
QY 570 lGluLysLeuIleLysAsnGlnLysGluSerHisPheValSerAlaArgProGlnSerGl 590
Db : : : : :
QY 1500 G---CAGTCTCCG 1509
Db : : : : :
QY 590 nSerGlnSerPro 594
Db : : : : :
RESULT 10
US-09-424-283-1
; Sequence 1, Application US/09424283
; Patent No. 6437219
; GENERAL INFORMATION:
; APPLICANT: Grimes, et al.
; TITLE OF INVENTION: Sucrose binding proteins
; FILE REFERENCE: 4630-50206
; CURRENT APPLICATION NUMBER: US/09/424,283
; CURRENT FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: PCT/US98/10465
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/047,568
; PRIOR FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Glycine max
US-09-424-283-1
Alignment Scores:
Pred. No.: 3,16e-07 Length: 524
Score: 167.00 Matches: 104
Percent Similarity: 34.0% Conservative: 89
Best Local Similarity: 18.3% Mismatches: 179
Query Match: 6.0% Indels: 196
DB: Gaps: 22
US-10-728-323-3 (1-1524) x US-09-424-283-1 (1-524)
QY 1 CGCAGCAACCCGAGGAGAAACGCGTCCAGTTCACCGCTCAATCGCAGACACCT--- 57
Db : : : : :
QY 92 ArgGluArgGluGluGlnGlnGluGlnHisGluGluGlnAsp---GluAsnProTyr 110
Db : : : : :
QY 58 -----GACAATCGCATTTGAATCAGAGGCGGTTCATATGAG 93
Db : : : : :
QY 111 lIlePheGluGluAspLysAspPheGluThrArgValGluThrGluGlyGlyArgIleArg 130
Db : : : : :
QY 94 ACTTGAACCCCAACCAACAGGAGTTCGAA---TGGCCCGCGGTCCGCCCTCTCTCGTTA 150
Db : : : : :
QY 131 ValLeuLysLysPheThrGluLysSerLysLeuLeuGlnGlyIleGluAsnPheArgLeu 150
Db : : : : :
QY 151 GTCCTC-----CGCCGCAACGCCCTTCGTAGGCTTCTACTCCATGCTCCCCAG 201
Db : : : : :
QY 151 AlaIleLeuGluAlaargAlaHisThrPheValSerProArgHisPheAspSerGluVal 170
Db : : : : :
QY 202 GAGATCTTCATCCAGCAAGGAGGATACTTTGGGTTCGATATTCCTGCTGCTCTAGA 261
Db : : : : :
```

```
Db 171 ValPhePheAsnIleLysGlyArgAlaValLeuGlyLeuVal----- 184
QY 262 CACTATGAAGAGCCTCACACAAAGTCTGCTGATCTCAGTCCCAAGACCAAGACGCT 321
Db 185 -----SerGluSerGluThr----- 189
QY 322 CTCCAAGGAAGAGACCAAGCCAAAGCAACAGCAAGAGATGTCACCAAGAGTGCACCGTTTC 381
Db 190 -----GluLysIle-----ThrLeu 194
QY 382 GATGAGGTGATCTCATTCAGTCTCCACCAGTGTGCTTCTGCTCTACCAACACAC 441
Db 195 GluProGlyAspMetIleHisProAlaGlyThrProLeuTyIleValAsnArgAsp 214
QY 442 GACACTGATGTTGTTGCTGTTCTCTTACTGACACCAACCAACAGCAACCAAGCTTGAT 501
Db 215 GluAsnAspLysLeuPheLeuAlaMetLeu----- 224
QY 502 CAGTTCCTCCAGGAGATTCAATTTGGCTGGGAACACGGAGCAAGATTTCTTAAGTACCAG 561
Db 224 ----- 224
QY 562 CAACAAGCAGACAAAGCAGACGAAGAGCTTACCATATAGCCCATACAGCCCGCAAGT 621
Db 225 -----HisIleProValSer-----ValSer 231
QY 622 CAGCCTAGACAAGAAGAGCGTGAATTTAGCCCTCGAGACACACAGCGCGCAGAGAACA 681
Db 232 ThrProGlyLysPheGluLeuPheAlaProGlyGlyArg----- 245
QY 682 CGAGACAAGAAGAAGAAACGAAGGTGGAACATCTTCAGCGGTTTCACGCCGGAGTTC 741
Db 246 -----AspProGluSerValLeuSerAlaPheSerTrpAsnVal 258
QY 742 CTGGAAACAGCCTTCAGGTTACACACAGACAGATAGTGCRAAACTTAAGAGCGAGACC 801
Db 259 LeuGlnAlaAlaLeuGlnThrProLysGlyLys-----LeuGluAsnValPhe 274
QY 802 GAGAGTGAAGAAGAGAGGAGCCATTCTGACAGTG---AGGGAGGCTCAGAACTTTGAGC 858
Db 275 AspGlnGlnAsnGluGlySerIlePheArgIleSerArgGluGlnValArgAlaLeuAla 294
QY 859 CAGATGAAGAAGAGAGCGTCCGACGAAGAGAGGAATACGATGAAGATGAATATGAATAC 918
Db 295 ProThrLysLys-----SerSerTrpTrpProPhe 304
QY 919 GATGAAGAGGATAGA-----AGCGTGGCAGGGGAGCAGACGCG 957
Db 305 GlyGlyGluSerLysProGlnPheAsnIlePheSerLysArgProThrIleSerAsnGly 324
QY 958 AGGGGGAATGGTATTGAAGAGAGCGATCTGCACCGCAAGTCTGACCGCAAGTCTAAAAAGAACATTGGTAGA 1017
Db 325 TyrGlyArgLeuThrGlu----- 330
QY 1018 AACAGATCCCTTGACATCTACAACCCCTCAAGTGGTTCTACTCAAAACTGCCAACGATCTC 1077
Db 331 -----ValGlyProAspAspGluLysSerTrpLeuGlnArgLeu 344
QY 1078 AACCTTCTAATACTTAGTGGCTTGACCTAGTGTGAATATGGAATCTCTACAGGAAT 1137
Db 345 AsnLeuMetLeu-----ThrPheThrAsnIleThrGlnArg 356
QY 1138 GCATTGTTTGTGCTCACTACAAACCAACGACACATCATATATCGATTGAGGGA 1197
Db 357 SerMetSerThrIleHisTyrAsnSerHisAlaThrLysIleAlaLeuValIleAspGly 376
QY 1198 CGGGCTCAGCTCAAGTC-----GTGGACAGCAACGCGCAACAGAGTGTAC 1242
Db 377 ArgGlyHisLeuGlnIleSerCysProHisMetSerSerArgSerSerHisSerLysHis 396
QY 1243 GACGAG-----GAGCTTCAAGAGGGTCACTGCTT 1272
Db 397 AspLysSerSerProSerTyrHisArgIleSerSerAspLeuLysProGlyMetValPhe 416
```

```
QY 1273 GTGGTGCACAGAACTTCGCC---GTCGCTGGAAGTCCAGAGCGAGAACTTCGAATAC 1329
Db 417 ValValProProGlyHisProPheValThrIleAlaSerAsnLysGluAsnLeuMet 436
QY 1330 GTGGCATTCAAGACAGACTCAAGGCCAGCATAGCC---AACCTCGCGGTGAAAACCTCC 1386
Db 437 IleCysPheGluValAsnAlaArgAspAsnLysPheThrPheAlaGlyLysAspAsn 456
QY 1387 GTCATAGAT-----AACCTCGCGGAGGAGGTG 1413
Db 457 IleValSerSerLeuAspAsnValAlaLysGluLeuAlaPheAsnTyrProSerGluMet 476
QY 1414 GTTGCAAATTTCATATGCGCTCCAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1458
Db 477 ValAsnGlyValPheLeuLeuGlnArgPheLeuGluArgLysLeuIleGlyArgLeuTyr 496
QY 1459 -----AAGAACAAACCCCTTCAAGTTCCTTCTGTTCCACCGTCT 1497
Db 497 HisLeuProHisLysAspArgLysGluSerPhePheProPheGluLeuProArgGlu 516
QY 1498 CAGCAGTCTCCGAGGCGGTGCGCT 1521
Db 517 GluArgGlyArgArgAlaAspAla 524

RESULT 11
US-09-805-694B-4
; Sequence 4, Application US/09805694B
; Patent No. 6864362
; GENERAL INFORMATION:
; APPLICANT: Kinney, Anthony
; TITLE OF INVENTION: Hypoallergenic Transgenic Soybeans
; FILE REFERENCE: B1432 US NA
; CURRENT APPLICATION NUMBER: US/09/805,694B
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/189, 823
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Glycine max
US-09-805-694B-4

Alignment Scores:
Pred. No.: 5,73e-07 Length: 454
Score: 164.00 Matches: 90
Percent Similarity: 33.3% Conservative: 75
Best Local Similarity: 18.2% Mismatches: 166
Query Match: 5.9% Indels: 164
DB: Gaps: 17

US-10-728-323-3 (1-1524) x US-09-805-694B-4 (1-454)
QY 178 CCTTTCTACTCCAATGCTCCCGAGAGATCTTCATCCAGCAAGGAGGAGTACTTGGG 237
Db 64 ProGlnTyrLeuAspSerAsnLeuIleIlePheIleArgGlyGluAlaLysLeuGly 83
QY 238 TTGATATTCCCTGGTTGTCTTAGACATATGAAGAGCCTCACACAAGGTGCTGATCT 297
Db 84 PheIleTyr----- 86
QY 298 CAGTCCCAAGACCAACCAAGAGCTCTCCAAGGAGAGAACCAAGCAACAGCAAGAGAT 357
Db 87 -----Asp 87
QY 358 AGTCACCAAGAGTGCACCGTTTCGATGAGGTGATCTTCATTGTCAGTTCACCAAGGTGT 417
Db 88 AspGluLeuAlaGluArgArgLeuLysThrGlyAspLeuTyrMetIleProSerGlySer 107
QY 418 GCTTTCTGGCTCTCAACGACACGACACTGATGTGTGTGCTGTTTCTTCTTACTGACACC 477
Db 418 GCTTTCTGGCTCTCAACGACACGACACTGATGTGTGTGCTGTTTCTTCTTACTGACACC 477
```

Db 108 AlaPheTyrLeuValAsnIleGlyGluArgLeuHisValIleCysSerIleAsp 127
QY 478 AACAAACGACAAACAGCTTGCATCATGTTCCCGAGAGATTCAATTTGGCTGGGAACAG 537
Db 128 ProSerThrSerLeuGlyLeuGluThrPhe---GlnSerPheTyrIleGlyGlyAla 146
QY 538 GAGCAAGAGTTCTTAAGGTACCACCAACAAAGCAGACAAAGCAGACGAAGACCTTACCA 597
Db 146 ----- 146
QY 598 TATAGCCCATACAGCCGCAAGTCAGCTAGACAAAGAGCGTGAAATTTAGCCCTCGA 657
Db 146 ----- 146
QY 658 GGACAGCACAGCGCAGAGAACGAGCAGGACAGAAAGAAAGAAAGAGGTGAAACATC 717
Db 147 AsnSerHisSer-----Val 151
QY 718 TTCAGGGCTTCACGCGGAGTTCCTGGAACAGCCTCCAGGTTCCAGGTTCACGACACAGATA 777
Db 152 LeuSerGlyPheGluProAlaIleLeuGluThrAlaPheAsn---GluSerArgThrVal 170
QY 778 GTGCAAAACCTAAGAGCGCAGACCGAGAGTGAAGAGAGGAGGCCATTGTGACAGTGGG 837
Db 171 ValGluGluIle-----PheSerLysGluLeuAspGlyProIleMetPheValasp 187
QY 838 GGAGGC-----CTCAGAACTTTGAGCCCGACAGATAAAG 870
Db 188 AspSerHisAlaProSerLeuTrpThrLysPheLeuGlnLeuLysLysAspLysGlu 207
QY 871 AGACGTGCCGACGAAGAGGAATACGATGAAGATGAATATGATACGATGAAGAGGAT 930
Db 208 GlnGlnLeuLysMetMetGln---AspGlnGluGlu-----AspGluGluGlu 223
QY 931 AGAAGCGTGGCAGGGAAGCAGA-----GGCAGGGGGAATGT 969
Db 224 LysGlnThrSerArgSerTrpArgLysLeuLeuGluThrValPheGlyLysValAsnGlu 243
QY 970 ATTGAAGACGATGTCGCCGCAAGTGTGTAAGAAAGACATTTGGTGAAGACAGATCCCT 1029
Db 244 LysIleGluAsn-----LysAspThrAlaGlySerProAlaSerTyr 257
QY 1030 GACATCTCAACCTCAAGCTGTCTCACTCAAACTGCC----- 1068
Db 258 AsnLeuTyrAspAspLysLysAlaAspPheLysAsnAlaTyrGlyTrpSerLysAlaLeu 277
QY 1069 -----AACGATCTCAACCTCTTAATCTAGGTGG 1098
Db 278 HisGlyGlyGluTyrProProLeuSerGluProAspIleGlyValLeuValLys--- 296
QY 1099 CTTGGACCTAGTCTGAATATGGAATCTCTACAGGAATGCATTTGTCGCTCACTAC 1158
Db 297 -----LeuSerAlaGlySerMetLeuAlaProHisVal 307
QY 1159 AACACCAACGACACAGCATATATCGATTTCAGGGGAGCGGCTCAGCTGCAAGTCGTG 1218
Db 308 AsnProIleSerAspGluTyrThrIleValLeuSerGlyTyrGlyGluLeuHisIleGly 327
QY 1219 GACAGCAACGGCAACAGAGTGTACGACGAGAGCTTCAAGAGGCTCACGTGCTTGTGGTG 1278
Db 328 TyrProAsnGlySerLysAlaMetLysThrLysIleLysGlnGlyAspValPheValVal 347
QY 1279 CGACAGAACTTCGCGTCGTGGA---AAGTCCAGAGCGAGAACTTCGAATACGTGGCA 1335
Db 348 ProArgTyrPheProPheCysGlnValAlaSerArgAspGlyProLeuGluPhePheGly 367
QY 1336 TTCAAGACGACTCAAGCCCGACATAGCCAAAC---CTCCCGGTGAAAACCTCCGCATA 1392
Db 368 PheSerThrSerAlaArgLysAsnLysProGlnPheLeuIleGlyAlaAlaSerLeuLeu 387
QY 1393 GATAACCTCCGAGAGGTGGTTGCAAAATTCATATGCGCTCCAAAGGGAGCAGGCAAG 1452
Db 388 ArgThrLeuMetGlyProGluLeuSerAlaAlaPheGlyValSerGluAspThrLeuArg 407

QY 1453 CAGCTTAAGAACAAACAAC-----CCCTTAAGTTCCTCGTTCACCG 1494
Db 408 ArgAlaValAspAlaGlnHisGluAlaValIleLeuProSerAlaTrpAlaAlaProPro 427
QY 1495 TCT-----CAGCAGTCTCCGAGGGCTGTG 1518
Db 428 GluAsnAlaGlyLysLeuLysMetGluGluProAsnAlaIle 442
RESULT 12
US-09-323-195A-18
; Sequence 18, Application US/09323195A
; Patent No. 6462257
; GENERAL INFORMATION:
; APPLICANT: Pullman, Gerald
; APPLICANT: Cairney, John
; APPLICANT: Perrera, Ranjan
; TITLE OF INVENTION: VICILIN-LIKE SEED STORAGE PROTEIN GENE PROMOTER AND
; TITLE OF INVENTION: METHODS OF USING THE SAME
; FILE REFERENCE: IPST0009
; CURRENT APPLICATION NUMBER: US/09/323,195A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Picea glauca
US-09-323-195A-18
Alignment Scores:
Pred. No.: 1,39e-06 Length: 448
Score: 160.00 Matches: 103
Percent Similarity: 33.6% Conservatives: 70
Best Local Similarity: 20.0% Mismatches: 172
Query Match: 5.8% Indels: 170
DB: Gaps: 20
US-10-728-323-3 (1-1524) x US-09-323-195A-18 (1-448)
QY 4 CAGCAACCGAGAGAACCGCTGCCAGTTCCAGCGCTCAATGCCGAGACCTGACAAT 63
Db 46 GluGluArgGluGluAsnProTyrValPhe-----HisSerAspSerPheArgThr 62
QY 64 CGCATTAATGATCAGAGCGCGGTACATTGAGACTTGGAAACCCCAACACACGAGGTTCCGA 123
Db 63 ArgAlaSerSerGluAlaGlyGluIleArgAlaLeu---ProAsnPheGlyGluValSer 81
QY 124 -----TGCGCCGCGCTGCCCTCTCTCGCTTAGTC-----CTCCGCCCAACGCC 168
Db 82 GluLeuLeuGluGlyIleArgLysPheArgValThrCysIleGluMetLysProAsnThr 101
QY 169 CTTGTGAGCGCTTCTTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGGA 228
Db 102 ValMetLeuProHisTyrIleAspAlaThrTrpIleLeuTyrValThrArgGlyArgGly 121
QY 229 TACTTTGGGTGATATTCCTCGTGTCTCTAGACACTGAGAGGCTCACACACAGGT 288
Db 122 TyrIleAlaTyrVal----- 126
QY 289 CGTCGATCTCAGTCCCAAGAACCCACCAAGACGCTCTCAAGGAGAGAACCAAGCAACAG 348
Db 127 -----HisGlnAsnGluLeuValLysArgLysLeu----- 136
QY 349 CAACGAGATAGTACCAGAAAGGTGCACCGTTTCGATGAGGTGATCTCATTTGAGTTCCC 408
Db 137 -----GluGluGlyAspValPheGlyValPro 145
QY 409 ACCGTTGTTCTTCTGGCTTACACAGCAGACACGACATGATGTTGTGCTGTTCTCTT 468
Db 146 SerGlyHisThrPheTyrLeuValAsnAsnAspHisAsnThrLeuArgIleAla--- 164
QY 469 ACTGACACCAACAACAACGACCAACCAAGCTTGATTCCTCCAGGAGATTCAATTTGGCT 528

```
Db 164 ----- 164
QY 529 GGAACACGAGCAGAGTCTTTAAGGTACCAGCAACAAGACAGACAGCAAGA 588
Db 164 ----- 164
QY 589 AGCTTACCATATAGCCCATACAGCCGCAAGAGTCCAGCTAGACAAAGAGCGGTGAATTT 648
Db 165 ----- 165
QY 649 AGCCCTCGAGGACACAGCCGACAGACAGCAGGACAGAGCAAGAAACAGAGT 708
Db 181 TyrValAlaGly ----- 189
QY 709 GGAACATCTTACAGCGCTTACGCGGAGTTCTCGAACAAGCCTTCCAGGTGACGAC 768
Db 190 ---ThrValTyrSerAlaPheSerAspValLeuGluAlaAlaPheAsnThrAsnVal 208
QY 769 AGACAGATAGTCGCAAAACCTAAGAGCGCGAGACCGAGAGTGAAGAGAGGAGCCATTGTG 828
Db 209 GlnGlnLeuGluArgIlePheGlyGly ----- HisLysSerGlyValIleIle 224
QY 829 ACAGTCAGGGGGCCTCAGAATCTTGAGCCCGAGTAGAAGAGAGCGTCCGACGAGAA 888
Db 225 ----- HisAlaAsnGluGlu 229
QY 889 GAGGAATACGATGAATATGAATACGATGAAGAGGATAGAGCGCTGCGAGGGA 948
Db 230 ----- GlnIleArgGluMetCysArgLysArgGly ---Phe 240
QY 949 AGCAGAGCGAGGGGAATGTTATGAAGAGACGATCTGCACCGCAAGTCTAAAAGAAAC 1008
Db 241 SerAlaGlySerMetSerAlaProGluHis ----- ProLysProPhe 254
QY 1009 ATTGGTAGAACAGATCCCTGACATCTACACCCCTCAAGCTGCTTCACTCAAAAGTGC 1068
Db 255 AsnLeuArgAsnGlnLysProAspPheGluAsn ---GluAsnGlyArgPheThrIleAla 273
QY 1069 AACGATCTCAACCTTCTAATCTTAGTGTGGCTTGGACCTAGTGTCTGAATATGGAATCTC 1128
Db 274 GlyProLysAsnTyrProPheLeuAspAlaLeuAspValSerValGlyLeuAlaAspLeu 293
QY 1129 TACAGGAATGCATTTGTCGTCTACTACACCAACGACACAGCATCATATATCGA 1188
Db 294 AsnProGlySerMetThrAlaProSerLeuAsnSerLysSerThrSerIleGlyIleVal 313
QY 1189 TTGAGGGGAGGGCTCAGCTGCAGTCTGCTG ----- 1218
Db 314 ThrAsnGlyGluGlyArgIleGluMetAlaCysProHisLeuGlyGlnHisGlyTrpSer 333
QY 1219 -----GACAGCAACGGCAACAGAGTGTACGACGAGGAGCTT 1254
Db 334 SerProArgGluArgGlyAspGlnAspIleThrTyrGlnArgValTrp ---AlaLysLeu 352
QY 1255 CAAGAGGGTCAGCTGTGTGGTCCACAGAACTTCGCGTCTGCTGGA ---AAGTCCCCAG 1311
Db 353 ArgThrGlySerValTyrIleValProAlaGlyHisProIleThrGluIleAlaSerThr 372
QY 1312 AGCGAGAACTTCGATACCTGGCATTCAGACAGACTCAAGCCCGACGATACCCAAAC --- 1368
Db 373 AsnSerArgLeuGlnIleLeuTrpPheAspLeuAsnThrArgGlyAsnGluArgGlnPhe 392
QY 1369 CTCGCGGTGAAACCTCCGTCATAGATAACCTCGCGGAGGAGTGTGTTCAAAATTCATAT 1428
Db 393 LeuAlaGlyLysAsnAsnValLeuAsnThrLeuGluArgGluIle ----- 407
QY 1429 GGCCTCCAAAGGAGGAGCAAGGCAAGCTTAAGAACAAACACCC 1473
Db 408 -----ArgGlnLeuSerPheAsnValPro 415
RESULT 13
US-09-323-195A-17
```

```
; Sequence 17, Application US/09323195A
; Patent No. 6462257
; GENERAL INFORMATION:
; APPLICANT: Pullman, Gerald
; APPLICANT: Cairney, John
; APPLICANT: Perrera, Ranjan
; TITLE OF INVENTION: VICILIN-LIKE SEED STORAGE PROTEIN GENE PROMOTER AND
; FILE REFERENCE: IPST0009
; CURRENT APPLICATION NUMBER: US/09/323,195A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 523
; TYPE: PRN
; ORGANISM: Pinus taeda
US-09-323-195A-17

Alignment Scores:
Pred. No.: 2,35e-06 Length: 523
Score: 158.00 Matches: 107
Percent Similarity: 32.7% Conservative: 74
Best Local Similarity: 19.3% Mismatches: 190
Query Match: 5.7% Indels: 183
DB: 2 Gaps: 20

US-10-728-323-3 (1-1524) x US-09-323-195A-17 (1-523)
```

```
QY 10 CCGAGGAGAGAACCGCTGCGAGTTCAGCGCCTCAATGCGCAGACCTGACAAAT----- 63
Db 37 ProGluAspHisGlyArgGlyHisGlnArgArgGluGluGluArgGluAsnProTyr 56
QY 64 -----CGCATTTGAATCAGAGGGCGGTTACATTGAGACT 96
Db 57 ValPheHisSerAspArgPheArgMetArgAlaSerSerAspAlaGlyGluIleArgAla 76
QY 97 TGGAAACCCCAACCAACAGGAGTTCGAA-----TGC GCGCGGCTCGCCCTCTCTCGCTTA 150
Db 77 Leu---ProAsnPheGlyGluAlaSerGluLeuLeuGluGlyIleSerLysTyrArgVal 95
QY 151 GTC-----CTCGCGGCGAACCCCTTCTGAGGCTTTCTACTCCAACTGCTCCCCAG 201
Db 96 ThrCysIleGluMetArgProAsnThrValMetLeuProHisTyrLeuAspAlaThrTrp 115
QY 202 GAGATCTTCATCCAGCAAGGAGGATACTTTGGGTTGATATTCCTGGTTGCTCTAGA 261
Db 116 IleLeuTyrValThrGlyArgGlyTyrIleAlaTyrVal----- 129
QY 262 CACTATGAAGAGCCTCACACACAGGTGCTGATCTCAGTCCCAAGACACCAAGACGT 321
Db 130 -----HisGlnAsnGluLeuValLysArgLys 138
QY 322 CTCCAAGGAGAGAACCAAGCCACGACGATAGTCACCAAGAGTGCCACCGTTTC 381
Db 139 Leu----- 139
QY 382 GATGAGGTGATCTCATTCGAGTCTCCACCGGTGTTGCTTCTGCTCTACAACGACCAC 441
Db 140 GluGluGlyAspValPheGlyValProSerGlyHisThrPheTyrLeuValAsnAsnAsp 159
QY 442 GACACTGATGTTGTTGCTGTTTCTCTACTTGACACCAACCAACGACCAACCGCTTGTAT 501
Db 160 AspHisAsnSerLeuArgIleThr----- 167
QY 502 CAGTTCCCGCAGGAGATTCATTTGGCTGGGAACACGGAGCAAGAGTCTTAAAGGTACCAG 561
Db 168 -----SerLeuLeuArgThrVal 173
QY 562 CAACAAAGCAGACAAAGCAGACGAGAGAGCTTTACCATATAGCCCATACAGCCCGCAAGT 621
Db 174 SerThrMetArgGlyGlu-----TyrGluProTyr----- 183
```

Qy	622	CAGCCTAGACAAGAAGCGTGAATTATTCAGCCTCGAGGACAGCACGCCGCGAGAACGA	681
Db	184	-----TyrValAla	186
Qy	682	GCAGGCAACAAGAAGAAAACGAAGGTGGAAACATCTTCAGCGCTTCACGCCGAGTTC	741
Db	187	GlyGlyArgAsnProGlu-----ThrValTyrSerAlaPheSerAspAspVal	202
Qy	742	CTGGAACAACAGCCTTCACAGTTGTACACGACAGACAGATAGTGCACAAACCTTAAGAGCGCAGACC	801
Db	203	LeuGluAlaAlaPheAsnThrAsn-----	210
Qy	802	GAGAGTGAAGAAGAGGAGCCATTGTGCACAGTGAAGGAGCGCTCAGAAATCTTGAGGCCCA	861
Db	211	-----ValIleGluAlaArgThrHisPheProVal-----	220
Qy	862	GATAGAAAGAGACGTCCGACGAAGAAGAGGAATACGATGAAGATGAATATGAATACGAT	921
Db	221	-----HisIleGluArgGluSerTyrSerMetAla	230
Qy	922	GAAGAGGATAGAAGCGCTGGCAGGGGAAGCAGAGCGAGGGGAATGTTATGAGAGACG	981
Db	231	AsnGluGluGlnIleArg--GluMetLeuArgLysArgGlyPheSerAlaGluSerMet	249
Qy	982	ATCTGCACCGCAAGTGTCTAAAGAAACATTTGGTGAACACAGATCCCTCGATCATCTACAAC	1041
Db	250	SerAlaSerGluHisProLysProPheAsnLeuArgAsnGlnLysProAspPheGluAsn	269
Qy	1042	CCTCAAGCTGGTTCACTCAAAACTGCCAAACGATCTCAACCTCTTAATACCTTAGGTGGCTT	1101
Db	270	AspAsnGlyArgPheThrArgAlaGlyProAsnGluAsnProLeu--LeuAspAlaVal	288
Qy	1102	GGACCTAGTGTGTAATATGGAATCTCTACAGGAATGCATTTGTTGCTCTACTACAAC	1161
Db	289	AspValThrAlaGlyPheGlyValLeuAsnProGlyThrMetThrAlaProSerHisAsn	308
Qy	1162	ACCAACCCACACACATATATCGATTGAGGGGACGGGCTCACGTGCAAGTCGTG---	1218
Db	309	ThrLysAlaThrSerIleAlaIleValThrGlnGlyGluGlyArgIleGluMetAlaCys	328
Qy	1219	-----GACAGCAAC	1227
Db	329	ProHisLeuGlyGlnHisGlyTyrPheSerArgArgGluLysGlyAspGlnGluIleAsn	348
Qy	1228	GGCAACAGAGTGTACGAGGAGGCTTCACAGAGGTCACGTGTTGTTGGTCCACAGAAC	1287
Db	349	TyrGlnArgVal--ArgAlaArgLeuArgThrGlyThrValTyrValValProAlaGly	367
Qy	1288	TTCCGCTCGCTGGAAAGTCCACAGCGAG--AACTTCGAATACGTGGCATCAACACA	1344
Db	368	HisProIleThrGluIleAlaCysThrGluGlyHisLeuGluIleLeuTyrPheAspIle	387
Qy	1345	GACTCAAGGCCACCATAGGCCAAC---CTGCCGGTGAATACTCCGTCATAGATAACCTG	1401
Db	388	AsnThrSerGlyAsnGluArgGlnPheLeuAlaGlyLysTyrAsnValLeuGlnThrLeu	407
Qy	1402	CCGAGGAGGTTGTTGCMAATTATATGCGCTCCAAAGG-----GACGAG	1446
Db	408	GluLysGluValArgGlnIleSerPheAsnIleProArgGlyGluGluLeuAspGluVal	427
Qy	1447	GCAAGGCAGCTTAAGAAC-----	1470
Db	428	LeuArgArgGlnLysAspGlnValIleLeuArgGlyProGlnMetGlnArgArgAspGlu	447
Qy	1471	CCCTTCAAGTTCTTCGTTCCACCGTCTCAGCAGTCTCCGA	1510
Db	448	ProArgSerSerIleHisAlaIleIleAlaAlaArg	460

RESULT 14

RESUL I 4
US-09-106-872A-17

US-09-106-872A-1 /
: Sequence 17, Application US/09106872A

: Sequence 17, Application
: Patent No. 6486311

Patent No. 6486311


```
Db      428 SerIlePheValProHisTyrAsnSerLysAlaThrPheValValLeuValAsnGluGly 447
QY      1198 CGGGCTCACGTGCAAGTCGTGGAC----- 1221
      ::|||::|::|::|
Db      448 AsnGlyTyrValGluMetValSerProHisLeuProArgGlnSerSerPheGluGluGlu 467
QY      1222 -----AGCAACGGCAACAGAGTG 1239
      ::|||::|::|::|
Db      468 GluGluGlnGlnGlnGluGlnGluGluGluArgSerGlyGlnTyrArgLys 487
QY      1240 TACGACGAGGAGTTCAAGAGGGTCACGTGCTTGTGGTCCACAGAACTTCGCCGTCGCT 1299
      ::|||::|::|::|
Db      488 IleArgSerGlnLeuSerArgGlyAspIlePheValValProAlaAsnPheProValThr 507
QY      1300 ---GGAAAGTCCAGAGCGAGAACTTCGAATACGTGGCATTC----- 1338
      |||||::|::|::|
Db      508 PheValAlaSerGlnAsnGlnAsnLeuArgMetThrGlyPheGlyLeuTyrAsnGlnAsn 527
QY      1339 ---AAGACAGACTCAAGGCCCATAGCCAACTCGCCGTGAAACTCCGTCATAGAT 1395
      |||||::|::|::|
Db      528 IleAsnProAspHisAsnGlnArgIlePhe---ValAlaGlyLysIleAsnHisValArg 546
QY      1396 AACCTGCCGAGGAGGTGTTCAAATTATATGGCTCCAAAGGGAGGCAAGGCAG 1455
      ::|||::|::|::|
Db      547 GlnTrpAspSerGlnAlaLysGluLeuAlaPheGlyValSerSerArgLeuValAspGlu 566
QY      1456 CTTAAGAACAACACCCCTTCAAG---TTCTTCGTTCCACCGTCTCAGCAGTCTCCGAGG 1512
      ::|||::|::|::|
Db      567 IlePheAsnAsnAsnProGlnGluSerTyrPheVal-----SerArgGlnArgGlnArg 584
QY      1513 GCT 1515
Db      585 Ala 585
```

Search completed: May 15, 2006, 22:28:33
Job time : 64.8306 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 23:34:07 ; Search time 43.013 Seconds
(without alignments)
4441.255 Million cell updates/sec

Title: US-10-728-323-3

Perfect score: 2779

Sequence: 1 cggcagcaaccggaggagaa.....ctccagggtgtggcttaa 1524

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlp
-Q=/abs/ABSWEB_spool/US10728323/runat_15052006_172145_22607/app_query.fasta_1
-DB=Published Applications AA Main -OFMT=fastan -SUFFIX=n2p.rapbm
-MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=US10728323@CGN 1 1_503@runat_15052006_172145_22607 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications AA Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pcp.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2673	96.2	507	4	US-10-245-871-96 Sequence 96, Appl
2	2673	96.2	507	4	US-10-253-286-96 Sequence 96, Appl
3	2673	96.2	510	4	US-10-228-806-6 Sequence 6, Appl
4	2673	96.2	510	4	US-10-100-303A-90 Sequence 90, Appl
5	2638	94.9	526	3	US-09-731-221-79 Sequence 79, Appl
6	2621	94.3	530	5	US-10-899-551-6 Sequence 6, Appl
7	1665	59.9	351	3	US-10-899-551-58 Sequence 58, Appl
8	1502	54.0	481	3	US-09-759-967-21 Sequence 21, Appl
9	1502	54.0	481	4	US-10-424-599-171702 Sequence 171702,
10	1502	54.0	488	4	US-10-425-114-43865 Sequence 43865, A
11	1502	54.0	491	4	US-10-425-114-44047 Sequence 44047, A

12	1502	54.0	492	4	US-10-425-114-43836 Sequence 43836, A
13	1502	54.0	492	4	US-10-425-114-46294 Sequence 46294, A
14	1499.5	54.0	466	4	US-10-425-114-71959 Sequence 71959, A
15	1499.5	54.0	466	4	US-10-425-114-71973 Sequence 71973, A
16	1499.5	54.0	485	3	US-09-759-967-20 Sequence 20, Appl
17	1499.5	54.0	488	4	US-10-425-114-45645 Sequence 45645, A
18	1499.5	54.0	489	4	US-10-425-114-43983 Sequence 43983, A
19	1499.5	54.0	489	4	US-10-425-114-44676 Sequence 44676, A
20	1499.5	54.0	489	4	US-10-425-114-45612 Sequence 45612, A
21	1499.5	54.0	489	4	US-10-425-114-50117 Sequence 50117, A
22	1499.5	54.0	489	4	US-10-425-114-50207 Sequence 50207, A
23	1499.5	54.0	489	4	US-10-425-114-51412 Sequence 51412, A
24	1499.5	54.0	489	4	US-10-425-114-58429 Sequence 58429, A
25	1499.5	54.0	489	4	US-10-425-114-71902 Sequence 71902, A
26	1499.5	54.0	489	4	US-10-425-114-71909 Sequence 71909, A
27	1499.5	54.0	489	4	US-10-425-114-71957 Sequence 71957, A
28	1499.5	54.0	489	4	US-10-425-114-71958 Sequence 71958, A
29	1499.5	54.0	489	4	US-10-425-114-71967 Sequence 71967, A
30	1499.5	54.0	489	4	US-10-425-114-71995 Sequence 71995, A
31	1499.5	54.0	489	4	US-10-425-114-72287 Sequence 72287, A
32	1499.5	54.0	489	4	US-10-425-114-72288 Sequence 72288, A
33	1498	53.9	489	4	US-10-425-114-45639 Sequence 45639, A
34	1496	53.8	476	5	US-10-409-993-1 Sequence 1, Appl
35	1496	53.8	477	5	US-10-409-993-15 Sequence 15, Appl
36	1496	53.8	485	5	US-10-409-993-9 Sequence 9, Appl
37	1496	53.8	485	5	US-10-409-993-13 Sequence 13, Appl
38	1496	53.8	495	3	US-09-759-967-2 Sequence 2, Appl
39	1496	53.8	495	4	US-10-424-599-171705 Sequence 171705, A
40	1496	53.8	495	4	US-10-425-114-45626 Sequence 45626, A
41	1496	53.8	501	4	US-10-425-114-43904 Sequence 43904, A
42	1496	53.8	505	4	US-10-425-114-71930 Sequence 71930, A
43	1496	53.8	505	4	US-10-425-114-71962 Sequence 71962, A
44	1496	53.8	506	4	US-10-425-114-43901 Sequence 43901, A
45	1496	53.8	506	4	US-10-425-114-45624 Sequence 45624, A

ALIGNMENTS

RESULT 1

US-10-245-871-96
; Sequence 96, Application US/10245871
; Publication No. US2003023594A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: 11-KEY/ANTIGENIC EPIOTOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2013
; CURRENT APPLICATION NUMBER: US/10/245,871
; CURRENT FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-245-871-96

Alignment Scores:

Pred. No.: 1.95e-221 Length: 507
Score: 2673.00 Matches: 507
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 96.2% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-3 (1-1524) x US-10-245-871-96 (1-507)

QY 1 CGCGAGCAACCGGAGGAGAAACGGTCCAGTTCAGCGCTCAATGCCAGACCTGAC 60

|||||

Db 1 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 20
QY 61 AATCGCATTTGAATCAGAGCGCGTTTACATTGAGACTTGGAAACCCCAACACGAGAGTTTC 120
Db 21 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 40
QY 121 GAATGCGCGGGTCCCTCTCTCGCTTAGTCTCCGCGCGCAACGCCCTTCTGTAGGCTC 180
Db 41 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 60
QY 181 TTCTACTCCAATGCTCCCGAGAGATCTTCATCCAGCAAGGAGGGATACTTTGGTTG 240
Db 61 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 80
QY 241 ATATTCCCTGGTTGCTTAGACACTATGAAGAGCCTTCACACAAGGTCTGTCATCTCAG 300
Db 81 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgSerGln 100
QY 301 TCCCAAGACCCCAAGAGCTCTCCAAGGAGAGAGCAACCAAGCAACGAGATAGT 360
Db 101 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 120
QY 361 CACCAGAAGTGCACCGTTTCGATGAGGTGATCTCATTTGCAGTTCCACCGGTGTGCT 420
Db 121 HisGlnIleValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 140
QY 421 TTCTGGCTCTACCAACGACCACGACACTGATGTTGTTGCTGTTCTTACTGACACCAAC 480
Db 141 PheTrpLeuTyrAsnAspHisAspThrAspValAlaValSerLeuThrAspThrAsn 160
QY 481 AACAAAGCAGACAGCTTGATCAGTTCCCGAGAGATTCATTTGCTGGACACGGAG 540
Db 161 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 180
QY 541 CAAGAGTTCTTAAGGTACCAGCAACAAAGCAGACAAAGCAGACAGCAAGAGCTTACCATAT 600
Db 181 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 200
QY 601 AGCCCATACAGCCCGCAAGTTCAGCTTAGCAAGAGAGCGTGAATTTAGCCCTCGAGGA 660
Db 201 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 220
QY 661 CACACAGCCGAGAGACAGCAGCAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 221 GlnHisSerArgArgGluArgAlaGlyGlnGlnGluGluAsnGluGlyGlyAsnIlePhe 240
QY 721 AGCGGCTTACGCGGAGTTCTTGGAAACAGCTTCAGGTTCCAGCAGACACAGATAGTG 780
Db 241 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 260
QY 781 CAAACCTAAGAGCGGAGACCGAGAGTGAAGAGAGAGGAGCCATTGTGACAGTGAGGGGA 840
Db 261 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 280
QY 841 GGCTCAGATCTTGAGCCAGATAGAAAGACGCTGCCAGCAAGAGAGAGAGATACGAT 900
Db 281 GlyLeuArgIleLeuSerProAspArgLysArgAlaAspGluGluGluGluTyrAsp 300
QY 901 GAAGATCAATATGAATACGATGAAGCATAGAGGCGTGGCAGGGGAAACAGAGGCAGG 960
Db 301 GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg 320
QY 961 GGGAAATGGTATTGAAGAGAGATCTGCACCGCAAGTCTAAAGAACATTTGTTAGAAC 1020
Db 321 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 340
QY 1021 AGATCCCTGCATCTACACCTCTCAAGCTGTTCTCACTCAAACTGCCCAACGATCTCAAC 1080
Db 341 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 360
QY 1081 CTTTCTAATACTAGGTGGCTTGGACCTTAGTGTGAATATATGAATCTCTACAGGAATGCA 1140
Db 361 LeuLeuIleLeuArgTyrLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 380

QY 1141 TTGTTTTCGCTCACTACAACACCAACGACACATCATATATCGATTGAGGGACGG 1200
Db 381 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 400
QY 1201 GCTCACGTGCAAGTCGTGACAGCAACCGCAACAGAGTGTACGACGAGAGGCTTCAAGAG 1260
Db 401 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 420
QY 1261 GGTACGCTGTTGTCGCCACAGAACTTCGCCGTGCTGGAAAGTTCAGAGCGAGAAC 1320
Db 421 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 440
QY 1321 TTCGAATAGTCGCTCAATCAAGACACACTCAAGGCCAGCATAGCCAACTCCGCCGTGAA 1380
Db 441 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 460
QY 1381 AACTCCGTCATAGATAACCTGCCGAGAGGTGTTGCAAAATTCATATGGCCTCCAAAGG 1440
Db 461 AsnSerValIleAspAsnLeuProGluGluValAlaAlaAsnSerTyrGlyLeuGlnArg 480
QY 1441 GAGCAGGCAAGCAGCTTAAGAACCAACCCCTTCAAGTTCCTCGTCCACCGTCTCAG 1500
Db 481 GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPheValProProSerGln 500
QY 1501 CAGTCTCCGAGGCTGTGCT 1521
Db 501 GlnSerProArgAlaValAla 507
RESULT 2
US-10-253-286-96
; Sequence 96, Application US/10253286
; Publication No. US20040058881A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; TITLE OF INVENTION: Ii-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2015
; CURRENT APPLICATION NUMBER: US/10/253,286
; PRIOR FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 96
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-253-286-96
Alignment Scores:
Pred. No.: 1,95e-221 Length: 507
Score: 2673.00 Matches: 507
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 96.2% Indels: 0
DB: 4 Gaps: 0
US-10-728-323-3 (1-1524) x US-10-253-286-96 (1-507)
QY 1 CGGCACAACCGGAGGAGAACCGTCGACGCTCAATGCGCAGACACCTGAC 60
Db 1 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 20
QY 61 AATCGCATTTGAATCAGAGCGCGTTTACATTGAGACTTGGAAACCCCAACACGAGAGTTTC 120
Db 21 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 40
QY 121 GAATGCGCGCGGTCCCTCTCTCGCTTAGTCTCCGCGCGCAACGCCCTTCTGTAGGCTC 180
Db 41 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 60

QY 361 CACCAGAGGTGCACCGTTTCGATGAGGTGATCTCATTCGAGTTCCACCGGTGTTGCT 420
Db 124 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143
QY 421 TTCTGGCTCTACAAACACACGACACTGATGTTGTTGCTGTTCTCTTACTGACACCAAC 480
Db 144 PheTrpLeuTyraAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 163
QY 481 AACACGACAACAGCTTGATCAGTTCCCGAGAGATTCAATTTGGCTGGGACACGGAG 540
Db 164 AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnThrGlu 183
QY 541 CAGAGTTCTTAAGTTACAGCAACAAACGACAAAGCAGACGAGAGAGCTTACCATAT 600
Db 184 GlnGluPheLeuArgTyraGlnGlnSerArgGlnSerArgArgSerLeuProTyra 203
QY 601 AGCCCATACAGCCCGCAAGTTCAGCTCAGTCCCGAGAGAGCGTGAATTTAGCCCTCGAGGA 660
Db 204 SerProTyraSerProGlnSerGlnProArgGlnGlnGluArgGluPheSerProArgGly 223
QY 661 CAGCACGCGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 720
Db 224 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluGluGluGluGluGluGluGluGlu 243
QY 721 AGCGGCTTCACGCGGAGTTCCTGGAAACAGCCCTCCAGGTTGACGACGACGACGACGACG 780
Db 244 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 263
QY 781 CAAACCTAAGAGCGCAGACCGAGAGTGAAGAGAGAGGAGCCATTGTGACAGTGAGGGGA 840
Db 264 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGluGluGluGluGluGluGluGluGlu 283
QY 841 GGCCTCAGAACTTGAGCCCGACATAGAAAGAGACGTGCCGACGAAAGAGGAAATACGAT 900
Db 284 GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluTyraAsp 303
QY 901 GAGATCAATATCAATACGATGAGGATGAGGATGAGGCGTGGCAGGGGACGACGACGAG 960
Db 304 GluAspGluTyraGluTyraAspGluGluAspArgArgGlyArgGlySerArgGlyArg 323
QY 961 GCGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGCTAAAAAGAACATTTGTAGAAAC 1020
Db 324 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 343
QY 1021 AGATCCCTTGACATCTACAAACCTCAAGCTGTTCAATCAAACTGCAACGATCTCAAC 1080
Db 344 ArgSerProAspIleTyraAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 363
QY 1081 CTTCTAATCTTAGGTGGCTTGACCTAGTGTGCTGAATATGGAATCTCTACAGGAATGCA 1140
Db 364 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyraGlyAsnLeuTyraArgAsnAla 383
QY 1141 TTGTTTGTCTCACTACAAACACGACACGACATCATATATGATTTGAGGGGACGG 1200
Db 384 LeuPheValAlaHisTyraAsnThrAsnAlaHisSerIleIleTyraArgLeuArgGlyArg 403
QY 1201 GCTCAGCTGCAAGTGTGGGACGACGCAACAGAGTGACACGAGGAGCTTCAAGAG 1260
Db 404 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyraAspGluGluLeuGlnGlu 423
QY 1261 GGTACGCTGCTTGGTGCCACAGAACTTCGCCGTGCTGGAAGTCCACGAGCGAGAAC 1320
Db 424 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 443
QY 1321 TTCGAATACGTGGCATTCAGACAGACTCAAGGCCCATAGCAACCTTCGCCGTGGA 1380
Db 444 PheGluTyraValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 463
QY 1381 AACTCCGCTCATAGATAACCTGCCGAGGAGGTGGTTGCCAAATTCATATGCTCCAAAGG 1440
Db 464 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyraGlyLeuGlnArg 483
QY 1441 GAGCAGGCAAGGCAGCTTAAGAACAAACACCCCTTCAAGTTCTTCGTTCCACCGTCTCAG 1500

Db 484 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProSerGln 503
QY 1501 CAGTCTCCGAGGCTGTGGCT 1521
Db 504 GlnSerProArgAlaValAla 510
RESULT 4
US-10-100-303A-90
; Sequence 90, Application US/10100303A
; Publication No. US20030202980A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0166
; CURRENT APPLICATION NUMBER: US/10/100,303A
; CURRENT FILING DATE: 2002-03-18
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 90
; TYPE: PRT
; LENGTH: 510
; ORGANISM: Atachis hypogaea, Prot/Nucleo Ara h 3
US-10-100-303A-90
Alignment Scores:
Pred. No.: 1,95e-221 Length: 510
Score: 2673.00 Matches: 507
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 96.2% Indels: 0
DB: 4 Gaps: 0
US-10-728-323-3 (1-1524) x US-10-100-303A-90 (1-510)
QY 1 CGSGACAAACCGAGGAGAAACGCTGCCAGTTCCAGCGCTCAATCGCAGACAGCTGAC 60
Db 4 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 23
QY 61 AATCGCATTTGAATCAGAGGCGGTTCATTGAGACTTGGAACCCCAACACAGGAGTTC 120
Db 24 AsnArgIleGluSerGluGlyGlyTyraIleGluThrTrpAsnProAsnAsnGlnGluPhe 43
QY 121 GAATGCGCGGCGTCCCTCTCTCGCTTAGCTTACCTCCGCGCAACGCCCTTCGTAGCCCT 180
Db 44 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 63
QY 181 TTCTACTCCAATGCTCCCGAGAGATCTTCATCCAGCAAGGAGGATACTTTGGGTTG 240
Db 64 PheTyraSerAsnAlaProGlnGluPheIleGlnGlnGlyArgGlyTyraPheGlyLeu 83
QY 241 ATATTCCTCTGTTGTCCTAGACACTATGAAGACCTTCACACAAAGTTCGTCATCTCAG 300
Db 84 IlePheProGlyCysProArgHisTyraGluProHisThrGlnGlyArgArgSerGln 103
QY 301 TCCCAAGACCAACAGACGCTCTCCAGGAGAGACCAAGCCCAACAGCAACAGATAGT 360
Db 104 SerGlnArgProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 123
QY 361 CACCAGAAGGTGCACCGTTTCGATGAGGTGATCTCATTCGAGTTCCACCGGTGTTGCT 420
Db 124 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143
QY 421 TTCTGGCTCTACAAACACCAACGACACTGATGTTGTTGCTGTTCTCTTACTGACACCAAC 480
Db 144 PheTrpLeuTyraAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 163
QY 481 AACCAACGACACAGCTTGATCAGTTCCCGAGAGATTCAATTTGGCTGGGACACCGAG 540
Db 164 AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnThrGlu 183
QY 541 CAAGAGTTCTTAAGGTACCGACCAACAAAGCAGACAAAGCAGACAAAGAGAGCTTACCATAT 600

184 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 203
601 AGCCCATACAGCCGCAAGTTCAGCTAGACAAAGAGAGCGTGAATTTAGCCCTCGAGGA 660
204 SerProTyrSerProGlnSerGlnProArgGlnGlnGluArgGluPheSerProArgGly 223
661 CAGCAGCCGAG 720
224 GlnHisSerArgArgGluArgAlaGlyGlnGlnGluGluGluGluGluGluGluGluGlu 243
721 AGCGGCTTCACCGCGAGTTCCTGGACCAAGCTTCCAGCTTCAGCAGCAGCAGTAGTG 780
244 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 263
781 CAAACCTTAAGAGCGGAG 840
264 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGluGluGluGluGluGluGluGlu 283
841 GGCCTCAGAAATCTTGAAGCCAGATAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
284 GlyLeuArgIleLeuSerProAspArgLysArgAlaAspGluGluGluGluGluGluGlu 303
901 GAAGATCAATATGATACATACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
304 GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg 323
961 GCGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGTAAAGAGAGAGAGAGAGAGAGAG 1020
324 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 343
1021 AGATCCCTCGATCTACAACTCAAGCTGCTTCACTCAAACTGCCAAGCATCTCAAC 1080
344 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 363
1081 CTTCTAAATCTTAGTGGCTTGGACCTAGTCTGTAATATGGAATCTTACAGGAATGCA 1140
364 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 383
1141 TTGTTTGTGCTCCTACCAACACCAACGACACAGCATCATATATCGATTGAGGGGACGG 1200
384 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 403
1201 GTCACGTCAAGTCGTGACAGCAACCGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
404 AlaHisValGlnValValAspSerAsnGlyAsnAsgValTyrAspGluGluLeuGlnGlu 423
1261 GGTACGTCTTGTGGTCCACAGAACTTCGCGCTGCTGGAAAGTCCGACGAGCAGAGAAC 1320
424 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 443
1321 TTCGAATACGTGCTTCAAGACAGACTCAAGGCCAGCATAGCCCAACCTCGCCGCTGAA 1380
444 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 463
1381 AACTCCGTATAGATAAATCTCGGAGGAGGTGGTTGCAAAATTCATATGCGCTCCAAAGG 1440
464 AsnSerValIleAspAsnLeuProGluGluValAlaAsnSerTyrGlyLeuGlnArg 483
1441 GAGCAGGCAAGGAGCTTAAGAACCAACACCCCTTCAAGTCTTCTGTTCCACCGCTCAG 1500
484 GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPhePheValProProSerGln 503
1501 CAGTCTCCGAGGCTGTGGCT 1521
504 GlnSerProArgAlaValAla 510

RESULT 5
US-09-731-221-79
; Sequence 79, Application US/09731221
; Patent No. US20020018778A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael

; TITLE OF INVENTION: Passive Desensitization
; FILE REFERENCE: 2002834-0103
; CURRENT APPLICATION NUMBER: US/09/731,221
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Arachis
; OTHER INFORMATION: Hypogaea
US-09-731-221-79

Alignment Scores:
Pred. No.: 2,07e-218 Length: 526
Score: 2638.00 Matches: 502
Percent Similarity: 99.0% Conservative: 0
Best Local Similarity: 99.0% Mismatches: 5
Query Match: 94.9% Indels: 0
DB: 3 Gaps: 0

US-10-728-323-3 (1-1524) x US-09-731-221-79 (1-526)

QY 1 CGGCAGCAACCGGAGGAGAGAGCGGTGCCAGTTCAGCGCCTCAATGCCGAGAGACCTGAC 60
DB 5 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 24
QY 61 AATCGCATTAATCAGAGGGGGTACATTGAGACTTGGAAACCCCAACACCAGAGGTTTC 120
DB 25 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnAlaAsnAsnGlnGluPhe 44
QY 121 GAATCGCGCGCGTCCGCTCTCTCGCTTAGTCTCCGCCGCAACGCCCTTCGTAGGCCT 180
DB 45 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro 64
QY 181 TTCTACTCCAATGTCTCCAGGAGATCTTCATCCAGCAAGGAGGGGATACCTTGGGTTG 240
DB 65 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnArgGlyTyrPheGlyLeu 84
QY 241 ATATTCCCTGTTGCTTAGACACTATGAAGAGCCTCACACAGGTCGTGCATCTCAG 300
DB 85 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 104
QY 301 TCCCAAGACCAACCAAGACGCTCTCAAGCAGAGAGCAAAAGCAACAGCAACAGATAGT 360
DB 105 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 124
QY 361 CACCAAGAGGTGCACCGGTTTCGATGAGGGTGATCTCATTTGCAGTTCCACCGGTGTGCT 420
DB 125 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 144
QY 421 TCTGGCTCTACAGACCCAGCAGACTGATGTTGTGTGTTCTTCTTCTTACTGACACCAAC 480
DB 145 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 164
QY 481 AACCAACCAACCGCTTGCATCAGTTCCCGCAGGAGATTCAATTTCCGCTGGGAAACCGAG 540
DB 165 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 184
QY 541 CAAGAGTTCTTAAAGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600
DB 185 GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgArgSerLeuProTyr 204
QY 601 AGCCCATACAGCCCGCAAGTTCAGCCTAGACAGAGAGCGTGAATTTAGCCCTCGAGGA 660
DB 205 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 224
QY 661 CAGCAGCCCGCAG 720
DB 225 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluGluGluGluGluGluGlu 244

```
QY 721 AGCGGCTTACCGCCGAGTTCCTGGAAACAGCGCTTCCAGGTGACACACAGACAGATAGTG 780
Db |||||||
QY 245 SerGlyPheThrProGluAlaLeuPheGlnAlaPheGlnValAspArgGlnIleVal 264
Db |||||||
QY 781 CAAACCTTAAGAGCCGAGACCCGAGAGTGAAGAGAGGAGGCCATTGTGACGTAGGGGA 840
Db |||||||
QY 265 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 284
Db |||||||
QY 841 GGCCTCAGAACTTGTAGCCAGATAGAAAGAGACGTGCCGACGAAGAAGAGGAATACGAT 900
Db |||||||
QY 285 GlyLeuArgAlaLeuSerProAspArgLysArgAlaAspGluGluGluTyrAsp 304
Db |||||||
QY 901 GAAGATGAATATCAATACGATGAAGAGGATAGAGCGGTGGCAGGGGAAGACGAGCAGG 960
Db |||||||
QY 305 GluAspGluTyrAlaTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg 324
Db |||||||
QY 961 GGAATAGGTATTGAAGAGACGATCTGCACCGCAAGTGTAAAGAAACATTTGGTAGAAAC 1020
Db |||||||
QY 325 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 344
Db |||||||
QY 1021 AGATCCCTGACATCTACAAACCTCAAGCTGCTGCTCACTCAAACTGCCAACGATCTCAAC 1080
Db |||||||
QY 345 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 364
Db |||||||
QY 1081 CTTCTAATCTTAGGTGGCTTGACCTAGTGTGTAATATGGAATCTCTACAGGAATGCA 1140
Db |||||||
QY 365 LeuLeuIleLeuArgIlePheGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 384
Db |||||||
QY 1141 TTCTTTTGTGCTCACTACAAACACCAACCGCACACAGCATCATATATCGATTGAGGGGACGG 1200
Db |||||||
QY 385 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 404
Db |||||||
QY 1201 GCTCAGCTGCAAGTCTGTGGACACAAACGGCAACAGAGTGTACACAGAGAGCTTCAAGAG 1260
Db |||||||
QY 405 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 424
Db |||||||
QY 1261 GGTCAAGTCTGTGTGGTCACAGAACTTCGCGCTCGTGGAAAGTCCACAGACGAGAAC 1320
Db |||||||
QY 425 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 444
Db |||||||
QY 1321 TTCGAATACGTGCATTCAAGACAGACTCAAGGCCCCAGCATAGCCAACTCGCGGTGAA 1380
Db |||||||
QY 445 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 464
Db |||||||
QY 1381 AACTCGCTCATAGATAACCTGCGGAGAGAGGTGGTTCGCAATTCATATGCGCTCCAAAGG 1440
Db |||||||
QY 465 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 484
Db |||||||
QY 1441 GACGAGCAAGGAGCTTAAGAACAAACACCCCTTCAAGTTCTTCGTTCCACCGTCTCAG 1500
Db |||||||
QY 485 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProProSerGln 504
Db |||||||
QY 1501 CAGTCTCCGAGGCTGTGGCT 1521
Db |||||||
QY 505 GlnSerProArgAlaValAla 511
Db |||||||
RESULT 6
US-10-899-551-6
; Sequence 6, Application US/10899551
; Publication No. US20050063994A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Burks, A. Wesley
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Howard, Sosin B.
; APPLICANT: Bottomly, Kim H.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy
; FILE REFERENCE: 2002834-0233
; CURRENT APPLICATION NUMBER: US/10/899,551
; CURRENT FILING DATE: 2004-07-26
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
```

```
; LENGTH: 530
; TYPE: PRT
; ORGANISM: species Arachis hypogea
US-10-899-551-6

Alignment Scores:
Pred. No.: 6,08e-217 Length: 530
Score: 2621.00 Matches: 500
Percent Similarity: 98.6% Conservative: 0
Best Local Similarity: 98.6% Mismatches: 7
Query Match: 94.3% Indels: 0
DB: Gaps: 0

US-10-728-323-3 (1-1524) x US-10-899-551-6 (1-530)

QY 1 CGCGAGCAACCGGAGAGAACCGCTGCAGTTCAGCGCCTCAATCCGCGAGACACCTGAC 60
Db 24 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
QY 61 AATCCGATTGAATCAGAGGGCGGTTACATTGAGACTTGGAAACCCCAACCAACAGGAGTTTC 120
Db 44 AsnArgIleGluSerGluGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 63
QY 121 GAATGCGCGGCGTCCGCCCTCTCTCGCTTAGTCTCTCCGCCCAACGCCCTTCGTAGCCCT 180
Db 64 GluCysAlaGlyValAlaAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro 83
QY 181 TTCTACTCCAATGCTCCCGAGAGATCTTCATCCAGCAAGGAGGGGATACTTTGGTTG 240
Db 84 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 103
QY 241 ATATTCCCTGGTGTCTTAGACACATATGAAGAGCCTCACACACAAGTTCGTTCATCTCAG 300
Db 104 IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgArgSerGln 123
QY 301 TCCCAAGACCAACCAAGACGCTCTCCAAGGAGAGAACCAAAAGCCACACGACAGATAGT 360
Db 124 SerGlnArgProProIleArgLeuGlnGlyGluAspGlnSerGlnGlnGlnArgAspSer 143
QY 361 CACCAAGGTGCACCGTTTCATGAGGTGTATCTCATTCAGTTCCTCCACCGGTGTGCT 420
Db 144 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 163
QY 421 TTCTGGCTCTACAACACACACACACTGATGTGTGTGTCTCTCTCTACTGACACCAAC 480
Db 164 PheTrpLeuTyrAsnAspHisAspThrAspValAlaValSerLeuThrAspThrAsn 183
QY 481 AACACACACACACAGCTTCATCAGTTCCCGAGAGATTCAATTTGGCTGGGAACACGGAG 540
Db 184 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHisGlu 203
QY 541 CAAGAGTTCTTAAGTTACAGCAACAAAGCAGCAACAGCAGAGAGAGAGCTTACCATAT 600
Db 204 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgArgSerLeuProTyr 223
QY 601 AGCCCATACAGCCCGCAAAAGTCAGCTAGACAAAGAGAGCGTGAATTTGGCTCCGAGGA 660
Db 224 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 243
QY 661 CAGCACAGCCGAGAGACGAGCAGGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 244 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluGluGluGluGlyGlyAsnIlePhe 263
QY 721 AGCGGCTTACCGCGGAGTTCCTGGAACAAGCCCTCCAGGTTACACGACAGACAGATAGT 780
Db 264 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 283
QY 781 CAAACCTTAAGAGCCGAGACCCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 284 GlnAsnLeuArgGlyGluAsnGluSerGluGluGluGlyAlaIleValThrValArgGly 303
QY 841 GGCCTCAGAACTTGTAGCCCGAGATAGAAAGAGACGTGCCGAGAGAGAGAGAGAGAGAT 900
Db |||||||
```



```
Db 304 GlyLeuArgIleLeuSerProAspArgLysArgAlaAspGluGluGluTyrAsp 323
QY 901 GAAGATGAATATACATACAGAGGATAGAGCGTGGCAGGGAGAGCAGG 960
Db 324 GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlyArg 343
QY 961 GGGAAATGGTATTGAAGAGACGATTCGACCGCAAGTGTAAAGAAACATTCGTAGAAG 1020
Db 344 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 363
QY 1021 AGATCCCTCGACATCTACAACTCAAGCTCAAGTGTCTCACTCAAACTGCCAACCATCTCAAC 1080
Db 364 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 383
QY 1081 CTTCTAAATCTTAGTGGCTTGGACCTAGTGTCTGAATATGGAATCTCTACAGGAATGCA 1140
Db 384 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 403
QY 1141 TTGTTTGTCTGCTACTACCAACCAACGACGACAGCATCATATATCGATTGAGGGACGG 1200
Db 404 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 423
QY 1201 GCTCAGTCTCAAGTCTGACAGCAACGCGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
Db 424 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 443
QY 1261 GGTCACTGCTGTGTGGTGCACAGAACTTCGCGCTGCGTGGAAAGTCCACAGACGAGAAC 1320
Db 444 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 463
QY 1321 TTCGAATACGTGGCAATCAAGACAGACTCAAGCGCCAGCATAGCCAACTCCCGGTGAA 1380
Db 464 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 483
QY 1381 AACTCGCTCATGATTAACCTCGGAGGAGGTGTGCAATTCATATGGCCCTCCAAAGG 1440
Db 484 AsnSerValIleAspAsnLeuProGluGluValAlaAlaAsnSerTyrGlyLeuGlnArg 503
QY 1441 GAGCAGGCAAGGCGAGCTTAAGAACCAACACCCCTTCAAGTCTTCTGTTCCACCGTCTCAG 1500
Db 504 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProProSerGln 523
QY 1501 CAGTCTCCAGGCGGTGTGGCT 1521
Db 524 GlnSerProArgAlaValAla 530

RESULT 7
; Sequence 58, Application US/10899551
; Publication NO. US20050063994A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Burks, A. Wesley
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Howard, Sosin B.
; APPLICANT: Bottomly, Kim H.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy
; FILE REFERENCE: 2002834-0233
; CURRENT APPLICATION NUMBER: US/10/899,551
; CURRENT FILING DATE: 2004-07-26
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 58
; LENGTH: 351
; TYPE: PRT
; ORGANISM: species Arachis hypogea
US-10-899-551-58

Alignment Scores:
Pred. No.: 1 63e-134 Length: 351
Score: 1665.00 Matches: 315
Percent Similarity: 97.8% Conservative: 0
Best Local Similarity: 97.8% Mismatches: 7
```

```
Query Match: 59.9% Indels: 0
DB: 5 Gaps: 0
US-10-728-323-3 (1-1524) x US-10-899-551-58 (1-351)
QY 1 CGGAGCAACCGGAGAGAAACGCGTCCAGCTTCAGCGCTCAATGCGCAGAGACTGAC 60
Db 15 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 34
QY 61 AATCCGATTGAATCAGAGGCGGTTCATCTGAGACTTGGAACTCCCAACCAACAGAGGTTTC 120
Db 35 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 54
QY 121 GAATCGCGCGCGTCCGCTCTCTCGCTTAGTCTCCCGCAACGCGCTTCGTAGGCT 180
Db 55 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 74
QY 181 TTCTACTCAATGCTCCCGAGAGATCTTCATCGAGCAAGAGGGATACCTTTGGTTG 240
Db 75 PheTyrSerAsnAlaProGlnGlnIlePheIleGlnGlnArgGlyTyrPheGlyLeu 94
QY 241 ATATTCCCTGTTGTCTAGACACTATCAAGAGCTTCACACAAAGGTCGTGATCTCAG 300
Db 95 IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgArgSerGln 114
QY 301 TCCCAAGACCAACCAAGACGCTCTCAAGGAGAGACCAAGCCAAACAGCAACGAGATAGT 360
Db 115 SerGlnArgProProIleArgLeuGlnGlyGluAspGlnSerGlnGlnGlnArgAspSer 134
QY 361 CACCAAGAGGTGACCGTTTCGATGAGGTGATCTCATTTGCAAGTTCCACCGGTGTGCT 420
Db 135 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 154
QY 421 TTCTGGCTCTACACGACGACGACTGATCTGTTGCTGTTCTTCTTACTGACACCAAC 480
Db 155 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValAsnLeuThrAspThrAsn 174
QY 481 AACCAACCAACCAAGCTTGTGATTCAGTTCCTCCAGGAGATTCATTTTCGCTGGGAACACGAG 540
Db 175 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHisGlu 194
QY 541 CAAGAGTTCTTAAGTACAGCAACAAAGCAGACAAAGCAGACGAGCAAGAGCTTACCATAT 600
Db 195 GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgSerLeuProTyr 214
QY 601 AGCCCATACAGCGCGCAAGTCCAGCTAGACAGAGAGCGCTGAATTTAGCCCTCGAGA 660
Db 215 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 234
QY 661 CAGCACGCGCGCAGAGAACGAGCAGACCAAGAAAGAAAGAAAGAAAGTGGAAACATCTTC 720
Db 235 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluAsnGluGlyGlyAsnIlePhe 254
QY 721 AGCGGCTTCACGCGGAGTTCCTGGAACAAGCTTCAGGTTCCAGGTGACGACAGACAGATAGT 780
Db 255 SerGlyPheThrProGluPheLeuGlnAlaPheGlnValAspAspArgGlnIleVal 274
QY 781 CAABACCTTAAGAGCGGACCGAGAGTCAAGAGAGGAGCGCATTTGTCAGCTGAGGGA 840
Db 275 GlnAsnLeuArgGlyGluAsnGluSerGluGluGlyAlaIleValThrValArgGly 294
QY 841 GGCTTCAGAAATCTTGAGCCCGCAGATAGAAAGAGACGTCGCGACGAAAGAGAGAAATACCAT 900
Db 295 GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluGluTyrAsp 314
QY 901 GAAGATGAATATACATACAGAGGATAGAGCGCTGCGAGGGAGAGCAGGAGG 960
Db 315 GluAspGluTyrGluTyrAspGluGluAspArgArgArgGlyArgGlySerArgGlyArg 334
QY 961 GGGAAAT 966
Db 335 GlyAsn 336
```

RESULT 8

```

US-09-759-967-21
; Sequence 21, Application US/09759967
; Publication No. US20030166518A1
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Nebraska
; TITLE OF INVENTION: METHOD FOR ALLERGEN CHARACTERIZATION
; FILE REFERENCE: UNL 3001.01
; CURRENT APPLICATION NUMBER: US/09/759,967
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Glycine max
US-09-759-967-21

Alignment Scores:
Pred. No.:      2,12e-120      Length:      481
Score:          1502.00       Matches:     297
Percent Similarity: 71.6%    Conservative: 68
Best Local Similarity: 58.2% Mismatches:   91
Query Match:      54.0%     Indels:      54
DB:               3         Gaps:        7

US-10-728-323-3 (1-1524) x US-09-759-967-21 (1-481)
Qy   1   CGGCAGCAACCGGAGGAGAACGCGTCCAGCTTCAGCGCCTCAATGCGCAGACCTGAC 60
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db   23   ArgGluGlnProGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 42
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy   61   AATCGCATTCGAATCAGAGGGCGGTACATTAGACTTGGAAACCCCAACAACAGAGAGTTC 120
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db   43   AsnArgileGluSerGluGlyGlyPheileGluThrTrpAsnProAsnAsnLysProPhe 62
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy   121  GAATGCCGCGGTGCCTCTCTCGCTTAGTCTCTCGCGCAACGCCCTTCGTAGGCCT 180
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db   63   GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 82
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy   181  TTCTACTCCAATGCTCCCCAGGAGATCTTCATCAGCAAGGAAGGGGATACTTTGGGTTG 240
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db   83   SerTyThrAsnAlaproGlnIleuIleTyriIleGlnGlnSerGlyIlePheGlyMet 102
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy   241  ATATTCCTGGTTGTCCTAGACACTATGAAGAGCTCACACAAAGTGTCGATCTCAG 300
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db   103  IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 119
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy   301  TCCCAAAGACCACCAAGACGTCTCCAAGGAGAGAACCAAGCCAAACGACGAGATAGT 360
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db   120  SerSerArgPro-----GlnAspArg 126
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy   361  CACCAGAAGTGACCGTTTCGATCAGGTTGATCTCAATTGAGTTCACCGGTGTGCT 420
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db   127  HisGlnLysIleTyHisPheArgGluGlyAspLeuileAlaValProThrGlyPheAla 146
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy   421  TTCGTGCTCTACAGCAGCACACACACTGATGTTGTGTCTGTTTTCTTACTGACACCAAC 480
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db   147  TyrTrpMetTyraAsnGlnuAspThrProValAlaValSerLeuileAspThrAsn 166
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy   481  AACACAGCAACACGAGCTTCAGTTCCCAGGAGATTCAATTTGGCTGGGAACACGGAG 540
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db   167  SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrrLeuAlaGlyAsnGlnGlu 186
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy   541  CAAGAGTTCTTAAGTTACAGCAACAAAGCAGACAAAGCAGACGAAGGCTTACCATAT 600
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db   187  GlnGluPheLeuGlnTyrrGlnProGlnLysGlnGlnGly-----Gln 199
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy   601  AGCCCATACAGCCCGAAAGTCAGCTCAGCTAGACAGAGAGCGTGAATTAGCCCTCGGGA 660
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Db   200  -----Gly 200
      |||...|||...|||...|||...|||...|||...|||...|||...|||...|||...
Qy   661  CAGCAGCGCGGAGAGAACGACGACGAGGACAAAGAAAGAAAAACGAAGTGGAAAACATCTTC 720

```

; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 171702
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_126061C.1.pap
US-10-424-599-171702

Alignment Scores:
Pred. No.: 2,12e-120 Length: 481
Score: 1502.00 Matches: 297
Percent Similarity: 71.6% Conservative: 68
Best Local Similarity: 58.2% Mismatches: 91
Query Match: 54.0% Indels: 54
DB: Gaps: 7

US-10-728-323-3 (1-1524) x US-10-424-599-171702 (1-481)

```
QY 1 CGGCAGCAACCGGAGGAGAACGGGTGCCAGTTCACAGCGCTCAATCGCGAGACACTGAC 60
Db 23 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 42
QY 61 AATCGATTGAATCAGAGCGGTTACATTGAGACTTGGAACCCCAACACAGGAGTTC 120
Db 43 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 62
QY 121 GAATGCGCGGCTGCCCTCTCTCGTTAGTCTCCGCGCAACGCCCTTCGTAGCGCT 180
Db 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 82
QY 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGATCTTTGGTTG 240
Db 83 SerTyThrAsnAlaProGlnGluIleTyIleGlnGlnGlySerGlyIlePheGlyMet 102
QY 241 ATATTCCCTGGTTGCTTAGACACTATGAGAGCCTCACACAAGTGTGTCATCTCAG 300
Db 103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 119
QY 301 TCCCAAGACCAACCAAGACGCTCTCCAAGGAGAGAACCAAGCCCAACAGCAGATAGT 360
Db 120 SerSerArgPro-----GlnAspArg 126
QY 361 CACCAAGAGTGCACCGTTTCGATGAGGTGATCTCAATTCAGTTCACCGGTGTGCT 420
Db 127 HisGlnLysIleTyHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146
QY 421 TTCCTGCTCTAACACGACGACACTGATGTTGTTGTTGTTCTTCTTCTACTGACCAAC 480
Db 147 TyrTrpMetTyTrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 166
QY 481 AACCAACGACACACGAGTTCATGATTCCTCCAGGAGATTCATTTGCTGGGACACGGAG 540
Db 167 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyTrLeuAlaGlyAsnGlu 186
QY 541 CAAAGTCTTCAAGTACCAGCAACAAAGACGACAAAGACGACGACGACGACGACGACGAC 600
Db 187 GlnGluPheLeuGlnTyTrGlnProGlnLysGlnGlnGly----- 199
QY 601 AGCCCATACAGCCCGCAAGTCAAGCTAGACAAAGAGAGCGGTGAATTTAGCCCTCAGGA 660
Db 200 -----Gly 200
QY 661 CAGCAGCGCAGAGAACGAGCAGGACGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 201 ThrGlnSerGlnLysGlyLysArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
QY 721 AGCGGTTTCACGCGGAGTTCCTGGAAACCAAGCCTTCAGGTTGACACAGACAGATAGTG 780
Db 221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 239
QY 781 CAAACCTTAAGCGCAGACCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 840 -----Gly 840
```

```
Db 240 ArgLysLeuGlnGlyLeuAsnGluGluGluGluGlyValThrValLysGly 259
QY 841 GGCCTCAGAAATCTTGAGCCCA-----GATAGAAAGAGACGTCGCCGACGAAGAGAG 891
Db 260 GlyLeuSerValIleSerProProThrGluGluGlnGlnArgProGluGluGlu 279
QY 892 GAATACGATGAAGATGAATATGAATACGATGAAGAGGATAGAGGCGTGGCAGGGAAGC 951
Db 280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294
QY 952 AGAGCAGCGGGAATGTTTGAAGAGAGGATCTGCACCGCAAGTCTCTCTCTCTCTCTCT 1011
Db 295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311
QY 1012 GGTAGAAACAGATCCCTGACATCTACAAACCTCAAGCTGAGTTCCTCTCTCTCTCTCT 1071
Db 312 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrAlaThr 331
QY 1072 GATCTCAACCTTCTTAATCTTAGTGGTTCGACCTTAGTCTGAATATCTCTCTCTCT 1131
Db 332 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 351
QY 1132 AGGAATGCAATCTTCTGCTCCTACCTACACACCAAGCGCACAGCATCATATATCGATTG 1191
Db 352 LysAsnAlaMetPheValProHisTyTrAsnLeuAsnAlaAsnSerIleIleTyAlaLeu 371
QY 1192 AGGGCAGCGGCTCAGTGCAGTCTGGAGACAGCAACCGCAACAGAGTGTACGACGAGGAG 1251
Db 372 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 391
QY 1252 CTTCAAGAGGTCAGTGTCTGTGTGTCACAGAACTTCGCGCTCGCTGGAAAGTCCCAG 1311
Db 392 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 411
QY 1312 AGCGAGAACTTCGATATGTCGATTCAGACAGACTCAAGCGCCGACGATAGCCACCTC 1371
Db 412 SerAspAsnPheGluTyTrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431
QY 1372 GCGGTGAAACCTCCGTCATAGTAACTCCGCGAGAGGAGTGTGCAAAATTCATATGGC 1431
Db 432 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 451
QY 1432 CTTCAAGAGGACGACGACGAGGAGGAGTTCAGAAACCAACCCCTTCAAGTTCCTGCTTCCA 1491
Db 452 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 471
QY 1492 CCGTCTCAGCAGTCTCCGAGGCGCTGCGCT 1521
Db 472 ProLysGluSerGlnArgValAlaAla 481
```

RESULT 10

US-10-425-114-43865
; Sequence 43865, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 43865
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700645818_FLI.pap

US-10-425-114-43865

Alignment Scores:

Pred. No.: 2,13e-120 Length: 488
Score: 1502.00 Matches: 297
Percent Similarity: 71.6% Conservatives: 68
Best Local Similarity: 58.2% Mismatches: 91
Query Match: 54.0% Indels: 54
DB: 4 Gaps: 7

US-10-728-323-3 (1-1524) x US-10-425-114-43865 (1-488)

QY 1 CGCAGCAACCGGAGAGAACGGCTGCCAGTTCACGGCTCAATCGCGAGACCTGAC 60
DB 30 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 49
QY 61 AATCGGATTGAATCAGAGGGCGGTTCATTGAGACTTGGAAACCCCAACACAGGAGTTTC 120
DB 50 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 69
QY 121 GAATGCGCGGCTCGCCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGCCCT 180
DB 70 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 89
QY 181 TTCCTACTCCATCTCCCGAGGAGATCTTCATCCAGCAAGGAGGATCTTTGGGTG 240
DB 90 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 109
QY 241 ATATTCCCTGGTGTCTAGACACTATGAAGAGCCTCACACAAAGTCTCGATCTCAG 300
DB 110 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 126
QY 301 TCCCAAGACCAACCAAGACGTCTCCAAGGAGAGAACCAAGCAACAGCAACGAGATAGT 360
DB 127 SerSerArgPro-----GlnAspArg 133
QY 361 CACAGAGGTGACCGTTTCGATGAGGTGATCTATTCAGTTCACCGGTGTTGCT 420
DB 134 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 153
QY 421 TTCCTGGCTPACCAACACCAACGACTGATGTTGTTGCTGTTCTCTACTGACACCAAC 480
DB 154 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 173
QY 481 AACCAAGCAACACAGCTTGATCAAGTCCCGAGAGATTCAATTGGCTGGGAACACGGAG 540
DB 174 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 193
QY 541 CRAAGCTTCTTAAGGTACCAACCAACCAAGCAACCAAGCAACCAAGCAACCAAGCAAC 600
DB 194 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly----- 206
QY 601 AGCCCATACAGCCCGCAAAAGTACGCTTAGACGAGAGCGTGAATTTAGCCCTCGAGGA 660
DB 207 -----Gly 207
QY 661 CAGCACAGCCGACAGAACGAGCAGGACAGACAGAGAAACGAGGTGAACATCTTC 720
DB 208 ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluGluGlnGlyGlySerIleLeu 227
QY 721 AGCGGCTTACGCGGAGTTCTCGAACAGCCCTCCAGGTTGACGACACAGATAGTG 780
DB 228 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 246
QY 781 CAAACCTAAGAGCGCAGACGAGTGAAGAGAGGAGCCATTGTGACAGTGAAGGGA 840
DB 247 ArgLysLeuGlnGlyGluAsnGluGluGluGluGlyAlaIleValThrValLysGly 266
QY 841 GGCTCTAGATCTTTAGCCCA-----GATAGAAAGAGAGTGCCGACGAGAAAGAG 891
DB 267 GlyLeuSerValIleSerProThrGluGlnGlnGlnArgProGluGluGluGlu 286
QY 892 GAATACGATGAAGATGAATATGAATACGATGAAGAGGATAGAAGCGCTGCGCGGGAAGC 951

DB 287 LysProAsp-----CysaspGluLysAspLysHisCysGlnSerGlnSer 301
QY 952 AGAGCGAGGGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTCTTAAAGAACATTT 1011
DB 302 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 318
QY 1012 GGTAGAAACAGATCCCTCGACATCTACAAACCTCAAGCTGGTTCCTCAAACTGCCAAC 1071
DB 319 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrAlaThr 338
QY 1072 GATCTCAACCTTCTAATCTTAGTGGCTTGACCTAGTCTGAAATATCGAAATCTCTAC 1131
DB 339 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 358
QY 1132 AGGAATGCATTTGTCCTCCTACACACCAACGACACAGCATCATATATCGATTG 1191
DB 359 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 378
QY 1192 AGGGACGGGCTCAGTGCAGTGTGGACAGCAACGCGCAACAGAGTGTACGACGAGAG 1251
DB 379 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 398
QY 1252 CTTCAAGAGGGTCACGTGCTGTGTGTCACAGAACTTCGCGTCTCGTGGAAAGTCCCCAG 1311
DB 399 LeuGlnGluGlyValValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 418
QY 1312 AGCGAGAACTTCAATPACGTGCATTCAGACAGACTCAAGGCCACGATAGCAACCTC 1371
DB 419 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 438
QY 1372 GCGGTGAAACTCCGTCTATAGATACCTGCGGAGGAGGTGGTTCAAATTCATATGC 1431
DB 439 AlaGlyAlaAsnSerLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 458
QY 1432 CTCCTCAAGGAGCAGGCAAGGAGCTTAAGAAACAACCCCTTCAAGTTCCTTCGTTCCA 1491
DB 459 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 478
QY 1492 CCGTCTCAGCAGTCTCCGAGGGCTGTGCT 1521
DB 479 ProLysGluSerGlnArgArgValValAla 488
RESULT 11
US-10-425-114-44047
; Sequence 44047, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 44047
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700763253_FLI.pep
US-10-425-114-44047

Alignment Scores: 2,14e-120 Length: 491
Pred. No.: 1502.00 Matches: 297
Score: 71.6% Conservatives: 68
Percent Similarity: 58.2% Mismatches: 91
Best Local Similarity: 58.2%

Query Match:	54.0%	Indels:	54
DB:	4	Gaps:	7
US-10-728-323-3 (1-1524) x US-10-425-114-44047 (1-491)			
Qy	1	CGGCAGCAACCGGAGGAGACCGCTGCAGTTCAGCGCCTCAATGCGCAGACCTTGAC	60
Db	33	ArgGluInProGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp	52
Qy	61	AATCGCATTTGAATCAGAGGGCGGTTACATTGAGACTTGGAAACCCAAACACGAGGTTTC	120
Db	53	AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe	72
Qy	121	GAATGCGCGCGCTGCCTCTCTCGCTTAGTCTCTCGCCGCGCAACGCGCTTCGTAGCGCT	180
Db	73	GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro	92
Qy	181	TTCTACTCCAAATGCTCCCGAGGAGATCTTCAATCCAGCAAGGAGGGGATACTTTGGTGTG	240
Db	93	SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet	112
Qy	241	ATATTCCCTGGTGTCTCTAGACACTATGAAGAGCTCACACAAAGTGTCTGCATCTCAG	300
Db	113	IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln	129
Qy	301	TCCAAAGACACCAAGACGCTCTCCAGGAGAGAACCAAGCCAAACAGCAACGAGATAGT	360
Db	130	SerSerArgPro-----GlnAspArg	136
Qy	361	CACCAGAAAGTGACACCGTTTCGATGAGGTGATCTCATTTGCAGTTCACCGGTGTGCT	420
Db	137	HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla	156
Qy	421	TTCTGGCTCTACAAAGCACGACACTGATGTGTGTCTTCTCTTACTGACACCAAC	480
Db	157	TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn	176
Qy	481	AACAAACGACAAACAGCTTGATCAGTTCGCCAGGAGATTCATTTGGCTGGGAACACGG	540
Db	177	SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu	196
Qy	541	CAAGAGTTCTTAAGGTACCACCAACAAAGACAGACAAAGACAGACGAAGAAGCTTACC	600
Db	197	GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly-----	209
Qy	601	AGCCCATACAGCCCGCAAGTCAGCTAGACAGAGAGCGTGAATTTAGCCCTCGAGGA	660
Db	210	-----Gly	210
Qy	661	CAGCACAGCCGACAGAACGACGAGGACAGAAAGAAACGAGGTGGAAACATCTTC	720
Db	211	ThrGlnSerGlnLysGlyLysArgGlnGlnGluAsnGluGlyGlySerIleLeu	230
Qy	721	AGCGCTTCACGCGCGAGTTCCTCGAACACAGCTTCCAGGTTCAGCAGACACAGATAGT	780
Db	231	SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal	249
Qy	781	CAAAACCTTAGAGCGGACGACGAGAGTGAAGAGAGGAGGCCATTGTGACAGTGAAGGG	840
Db	250	ArgLysLeuGlnGlyGluAsnGluGluGluGlyGlyAlaIleValThrValLysGly	269
Qy	841	GGCTCTCAGAATCTTGACCCCA-----CATAGAAAGACCTGCCGACGAGAGAG	891
Db	270	GlyLeuSerValIleSerProProThrGluGlnGlnArgProGluGluGluGlu	289
Qy	892	GAATACATGAAGATGAATATGAATACGATCAAGAGGATGAAGGCGGTGGCAGGGGAAG	951
Db	290	LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer	304
Qy	952	AGAGCGAGGGGAATGTATTGAAGAGACGATCTGTCCACCGCAAGTGTCTAAAGAACATT	1011
Db	305	Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle	321

```
Db      34  ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 53
QY      61  AATCGATTGAATCAGAGCGGTTTACATTGAGACTTGGAAACCCCAACACAGGAGTTC 120
Db      54  AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 73
QY     121  GAATGCGCGCGCTCCCTCTCTCGCTTAGTCTCCTCGCGCGCAACGCCCTTCGTAGCCCT 180
Db      74  GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 93
QY     181  TTCTACTCAATCTCCCAAGAGATCTTCATCCAGCAAGGAAGGGGATACTTTGGTTG 240
Db      94  SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 113
QY     241  ATATTCCCTGGTTGCTTAGACACTATGAAGACCTTCACACAAGGTGCTGCATCTCAG 300
Db     114  IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 130
QY     301  TCCCAAGACCAACCAAGACGCTCTCCAAGGAGAGACCAAGCCAAACAGCAACGAGATAGT 360
Db     131  SerSerArgPro-----GlnAspArg 137
QY     361  CACCAAGAGGTGCACCGTTTCGATGAGGGTGATCTCATTCGAGTCTCCACCGGTGTGCT 420
Db     138  HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 157
QY     421  TTCCTGCTTACAAACGACCAACGACACTGATGTTGTGTTGTTCTTCTTACTGACCAAC 480
Db     158  TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 177
QY     481  AACAAACGACAAACAGCTTGATCAGTTCCTCCAGAGATTCATTTGGCTGGACACGGAG 540
Db     178  SerPheGlnAsnGlnLeuAspGlnMetProArgPheTyrLeuAlaGlyAsnGlnGlu 197
QY     541  CAAGAGTTCTTAAGGTACCAACAAAGCAGCAAAAGCAGACAGAGCAAGAGCTTTACCATAT 600
Db     198  GlnGluPheLeuGlnTyrGlnProGlnLysGlnGly----- 210
QY     601  AGCCCATACAGCCCGCAAGTCAGCTCAGCAAGAGAGCGTGAAATTAGCCCTCAGGA 660
Db     211  ----- 211
QY     661  CAGCACAGCGCAGAGAACGAGCAGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db     212  ThrGlnSerGlnLysGlyLysArgGlnGlnGlnGluGlnGlnGlnGlnGlnGlnGln 231
QY     721  AGCGGTTTCACGCGGAGTTCCTGGAAACAGCCTTCAGGTTGACGACACAGACAGATAGT 780
Db     232  SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal--AspArgGlnIleVal 250
QY     781  CAAAACTTAAGAGCGGACCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db     251  ArgLysLeuGlnGlyGluAsnGluGluGluGluGluGluGluGluGluGluGluGlu 270
QY     841  GGCCTCAGATCTTGAGCCCA-----GATAGAAAGAGACGTGCCGACGAGAGAG 891
Db     271  GlyLeuSerValIleSerProThrProThrGluGlnGlnGlnGlnGlnGlnGlnGln 290
QY     892  GAATACGATGAAGATGAATATGAATCAAGTCAAGAGGATAGAAGCGGTGCGAGGGGAAGC 951
Db     291  LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 305
QY     952  AGAGCGAGGGGAATGTTATTGAAGAGACGATCTGCACCCCAAGTGTAAAGAAACATT 1011
Db     306  Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 322
QY     1012  GGTAGAAAACAGATCCCTCGACATCTCAACACCTCAAGCTGTTTCACTCAAAACCTGCCAAC 1071
Db     323  GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr 342
QY     1072  GATCTCAACCTTCTAATCTAGGTGGCTTGGACCTAGTCTGAATATGAAATCTCTAC 1131
Db     343  SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 362
```

```
QY     1132 AGGAATGCATTTGTTTCTCGCTCACTACAACACACCAACGACACAGCATCATATATCATGTTG 1191
Db      363 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleTyrAlaLeu 382
QY     1192 AGGGAGCGGCTCACGTGCAAGTCGTGGACAGCAACGGCAACAGAGTGTACGACGAGGAG 1251
Db      383 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 402
QY     1252 CTTCAAGAGGGTCACGTGCTTGTGTGTCACAGAACTTCGCCCTCGCTCGTGAAGTCCCAG 1311
Db      403 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 422
QY     1312 AGCGAGAATCTCAATACGTGCATTCAGACAGACTCAAGCCCGACGATAGCAACCTC 1371
Db      423 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 442
QY     1372 GCGGGTGAACACTCCGTGCATAGATAACCTCGGAGGAGTGGTTGCAAAATTCATATGGC 1431
Db      443 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 462
QY     1432 CTCRAAGGAGCAGCAAGGAGGCTTAAGAACAACCAACCCCTTCAAGTTCTTCGTCCA 1491
Db      463 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 482
QY     1492 CGCTCTCAGCAGTCTCCGAGGCGCTGTGGCT 1521
Db      483 ProLysGluSerGlnArgArgValValAla 492
RESULT 13
US-10-425-114-46294
; Sequence 46294, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46294
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700757865_FLI.pep
US-10-425-114-46294
Alignment Scores:
Pred. No.: 2,14e-120 Length: 492
Score: 1502.00 Matches: 297
Percent Similarity: 71.6% Conservative: 68
Best Local Similarity: 58.2% Mismatches: 91
Query Match: 54.0% Indels: 54
DB: 4 Gaps: 7
US-10-728-323-3 (1-1524) x US-10-425-114-46294 (1-492)
```

```
QY      1  CGSAGCAACCGGAGGAGAACGCGTCGAGTTCCAGCGCTCAATCGCAGACAGCCTGAC 60
Db      34  ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 53
QY     61  AATCGATTGAATCAGAGCGGCTTACATTGAGACTTGGAAACCCCAACACAGGAGTTC 120
Db     54  AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 73
QY     121  GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCCTCGCGCAACGCCCTTCGTAGGCCCT 180
```

```
Db 74 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 93
:::|||||
181 TTCTACTCAATGCTCCAGAGATCTTCATCCAGCAAGGAGGGGATACCTTGGGTG 240
:::|||||
Db 94 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 113
:::|||||
Qy 241 ATATTCCTCGTGTCTAGACACTATGAAGACCTCACACACAAGGTGCTGATCTCAG 300
:::|||||
Db 114 IlePheProGlyCysProSerThrPheGluGluProGlnGlnGlyGly-----Gln 130
:::|||||
Qy 301 TCCAAAGACCAACAGAGCGTCTCAAGGAGAAGACCAAGCCACAGCAACAGAGATAGT 360
:::|||||
Db 131 SerSerArgPro-----GlnAspArg 137
:::|||||
Qy 361 CACCAGAAGGTGCACCGTTTCGATGAGGTGATCTCATTCAGTTCGCCACCGGTGTGCT 420
:::|||||
Db 138 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 157
:::|||||
Qy 421 TTCTGCTCTTACAAACGACACACACTGATGTTGTTGCTGTTCTTCTTACTGACACCAAC 480
:::|||||
Db 158 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 177
:::|||||
Qy 481 AACAAACGACACAGCTGATCAAGTTCCTCCAGAGATTCAAATTTGGCTGGGACACGGAG 540
:::|||||
Db 178 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 197
:::|||||
Qy 541 CAAGAGTCTTAAAGTACCAACCAACAGCAAGCAAGCAGACGAGAGCTTACCATAT 600
:::|||||
Db 198 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly----- 210
:::|||||
Qy 601 AGCCCATACAGCCGCAAAAGTCAGCCTAGACAAAGAGCGTCAATTATGACCTCGAGGA 660
:::|||||
Db 211 -----Gly 211
:::|||||
Qy 661 CAGCAGCGCGAGAGAACGAGCAGACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
:::|||||
Db 212 ThrGlnSerGlnLysGlyLysArgGlnGlnGluGlnGluLysGlnGlySerIleLeu 231
:::|||||
Qy 721 AGCGGCTTACGCGCGAGTCTCGGAACAAGCTTCAGGCTTCAGGTTGACGACAGATAGT 780
:::|||||
Db 232 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 250
:::|||||
Qy 781 CAAACCTAAGCGCGAGACCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
:::|||||
Db 251 ArgLysLeuGlnGlyLysGlnGluGluGluLysGlyAlaIleValThrValLysGly 270
:::|||||
Qy 841 GGCTCAGAACTTGTAGCCCA-----GATAGAAAGAGAGCGTCCGACCAAGAGAG 891
:::|||||
Db 271 GlyLeuSerValIleSerProThrGluGluGlnGlnGlnArgProGluGluGlu 290
:::|||||
Qy 892 GAATACGATCAAGATCAATATGATACGATGAGAGATAGAGCGTGGCGGAGGAGC 951
:::|||||
Db 291 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 305
:::|||||
Qy 952 AGAGGAGCGGGGAATGCTATTGAAGAGAGAGATCTGACCGCAAGTCTATAAAGAACATT 1011
:::|||||
Db 306 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisIle 322
:::|||||
Qy 1012 GGTAGAAAACAGATCCCTGACATCTCAACCTTCAGCTGGTTCTCACTAAAACCTGCCAAC 1071
:::|||||
Db 323 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr 342
:::|||||
Qy 1072 GATCTCAACTTCTATCTTAGTGGCTTGGACCTAGTCTCAATATGAAATCTCTAC 1131
:::|||||
Db 343 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 362
:::|||||
Qy 1132 AGGAATGCATGTTGCTCGCTCACTACAACACCAACGACACAGCATCATATATCATGTTG 1191
:::|||||
Db 363 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 382
:::|||||
Qy 1192 AGGGGACGGGCTCACTGCAAGTCTGTGGACAGCAACCGGCAACAGAGTGTACCAACGAGG 1251
:::|||||
```

```
Db 383 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 402
Qy 1252 CTTCAAGAGGGTCACGTCTGTGTGTGTCACACAGAACTTCGCGCTCGTGGAAAGTCCCAAG 1311
|||||
Db 403 LeuGlnGluGlyGlnValIleValProGlnAsnPheAlaValAlaAlaArgSerGln 422
|||||
Qy 1312 AGCGAAGAACTTCGAATACGTGGCATTAAGACAGACTCAAGGCCAGCATAGCCCAACCTC 1371
|||||
Db 423 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 442
|||||
Qy 1372 GCCGTGAAAACCTCGCTCATAGATAACCTGCCGAGGAGGTGGTTCGCAATTCATATGCG 1431
|||||
Db 443 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 462
|||||
Qy 1432 CTCCTCAAGAGGAGCAGGCAAGCAGCTTAAGAACAAACCCCTTCAAGTTCCTTCGTTCCA 1491
|||||
Db 463 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 482
|||||
Qy 1492 CCGTCTCAGCAGTCTCCGAGGCGCTGTGCT 1521
|||||
Db 483 ProLysGluSerGlnArgArgValValAla 492
|||||

RESULT 14
US-10-425-114-71959
; Sequence 71959, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screeen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 71959
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700676992_FLI.pap
US-10-425-114-71959
```

```
Alignment Scores:
Pred. No.: 3,45e-120 Length: 466
Score: 1499.50 Matches: 297
Percent Similarity: 72.0% Conservative: 71
Best Local Similarity: 58.1% Mismatches: 94
Query Match: 54.0% Indels: 49
DB: 4 Gaps: 6
```

US-10-728-323-3 (1-1524) x US-10-425-114-71959 (1-466)

```
Qy 1 CGGCAGCAACCGGAGGAGACGCGTCCAGTTCCAGCGCTCAATGCCAGAGACTGAC 60
|||||
Db 1 ArgGluGlnAlaGlnGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProAsp 20
|||||
Qy 61 AATCCCATTTGAATCAGAGGGCGGTTCACATTGAGACTTGGAACTTGAACCCCAACACCGAGGTTT 120
|||||
Db 21 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 40
|||||
Qy 121 GAATCGCGCGCGTCCGCTCTCTCGCTTAGTCTCTCCGCGCAACCGCCCTTCGTAGGCT 180
|||||
Db 41 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 60
|||||
Qy 181 TTCCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGGGATACCTTCGTTG 240
|||||
Db 61 SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyAsnGlyIlePheGlyMet 80
|||||
```



```
QY 361 CACCAGAGGTGCACCGTTTCGATGAGGTGATCTCATTCAGTTCCTCCACCGGTGTGCT 420
Db 108 HisGlnLysValHisArgPheArgGluGlyAspLeuIleAlaValProThrGlyValAla 127
QY 421 TTCTGGCTCTACAACGACGACGACACTGATGTGTGTGCTTCTTCTTACTGACACCAAC 480
Db 128 TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerIleIleAspThrAsn 147
QY 481 AACAACGACAAACAGCTTCAGTTCCTCCAGGAGATTCAAATTGCTGGGACACGAG 540
Db 148 SerLeuGluAsnGlnLeuAspGlnMetProArgPheTyrLeuAlaGlyAsnGlnGlu 167
QY 541 CAAGAGTTCCTTAAGGTACCAGCAACAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600
Db 168 GlnGluPheLeuLysTyrGlnGln-----GlnGlnGlyGlySerGlnSerGlnLysGly 176
QY 601 AGCCCATACAGCCCGCAAAAGTCAGCTAGACAAAGAGCGGTGAATTTAGCCCTCGAGA 660
Db 177 -----GlnGlnGlyGlySerGlnSerGlnLysGly 186
QY 661 CAGCAGCGCGCAGAGAACGAGCAGGACAGAGAGAGAGAAACGAAAGGTGGAACATCTTC 720
Db 187 LysGln-----GlnGluGluGluAsnGluGlySerAsnIleLeu 199
QY 721 AGCGGCTTCACGCGGAGTTCCTGGAACAAGCCTTCAGGTTTGACGACACAGACAGATAGTG 780
Db 200 SerGlyPheAlaProGluPheLeuLysGluAlaPheGlyVal---AsnMetGlnIleVal 218
QY 781 CAAACCTTAAGAGCGGAGACCGAGAGTGAAGAGAGGAGGCCATTTGTGACGTAGGGGA 840
Db 219 ArgAsnLeuGlnGlyGluAsnGluGluAspSerGlyAlaIleValThrValLysGly 238
QY 841 GCCTCAGAACTTGTAGCCAGACAGAGAGACGTGCCGACGAGAGAGAGGAATACGAT 900
Db 239 GlyLeuArgValThrAlaProAlaMetArgLysProGlnGlnGluAspAspAsp 258
QY 901 GAAGATGAA-----TATGAATACGATGAAGAGGATAGAAGCGGTGGCAGGGGA 948
Db 259 GluGluGlnProGlnCysValGluThrAspLysGlyCysGlnArgGlnSerLysArg 278
QY 949 AGCAGAGCGGGGGAATGGTATTGAAGAGAGACATCTGCCCGCAAGTCTAAAGAAC 1008
Db 279 SerArg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgGlnAsn 295
QY 1009 ATTGGTAGAAGACAGATCCCTCAGATCTACACCTCAAGCTGTTCTACTCAAACTGCC 1068
Db 296 IleGlyGlnAsnSerSerProAspIleTyrAsnProGlnAlaGlySerIleThrAla 315
QY 1069 AACGATCTCAACTTCTAATACCTTAGGTGGCTTGGACCTAGTGCTGAATATGAAATCTC 1128
Db 316 ThrSerLeuAspPheProAlaLeuTrpLeuLeuLysLeuSerAlaGlnTyrGlySerLeu 335
QY 1129 TACAGGAATGCAATTTGTCGTCTACTACACCAACGACGACAGCATCATATATCGA 1188
Db 336 ArgLysAsnAlaMetPheValProHisTyrThrLeuAsnAlaAsnSerIleIleTyrAla 355
QY 1189 TTGAGGGGACGGCTCACCTGCACTGGAGCAGCAACGACGACAGAGTGTACGACGAG 1248
Db 356 LeuAsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGly 375
QY 1249 GAGCTTCAAGAGGGTCACCGTGTGTGGTGCCACAGAACTTCGCGTCTCGTGGAAAGTCC 1308
Db 376 GluLeuGlnGluGlyGlyValLeuIleValProGlnAsnPheAlaValAlaLysSer 395
QY 1309 CAGAGCGGAATTCGAATACGTGGCATTCAGACAGACTCAAGGCCAGCATAGCCAAC 1368
Db 396 GlnSerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsn 415
QY 1369 CTCGCGGTGAAACTCCGTCATAGATAACCTCGCGGAGGAGGTGGTTCGAATTCATAT 1428
Db 416 LeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnHisThrPhe 435
```

```
QY 1429 GGCCTTCAAAGGAGCAGGCAAGGCGCTTAAGAACAAACCCCTTCAAGTTCCTCGTT 1488
Db 436 AsnLeuLysSerGlnGlnAlaArgGlnValLysAsnAsnProPheSerPheLeuVal 455
QY 1489 CCACGCTCTCAGCAGTCTCCGAGGGCTGTGCT 1521
Db 456 ProProGlnGluSerGlnArgArgAlaValAla 466
```

Search completed: May 16, 2006, 00:00:55
Job time : 240.065 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 23:41:32 ; Search time 5.24287 Seconds
(without alignments)
4094.130 Million cell updates/sec

Title: US-10-728-323-3
Perfect score: 2779
Sequence: 1 cggcagcaacggaggagaa.....ctccagggtgtggcttaa 1524

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 250354 seqs, 46948837 residues

Total number of hits satisfying chosen parameters: 500708

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlpl
-MODEL=frame+ n2p model
-Q=/abss/ABSSWEB_spool/US1078323/runat_15052006_172148_22650/app_query.fasta.1
-DB=Published Applications AA New -QFMT=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs802p
-USER=US1078323@cgn1_1_63@runat_15052006_172148_22650 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :					Published Applications_AA_New.*				
1: /SIDSS/ptodata/1/pubpaa/US08_NEW_PUB.pepl.*					2: /SIDSS/ptodata/1/pubpaa/US06_NEW_PUB.pep.*				
3: /SIDSS/ptodata/1/pubpaa/US07_NEW_PUB.pep.*					4: /SIDSS/ptodata/1/pubpaa/US08_NEW_PUB.pep.*				
5: /SIDSS/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*					6: /SIDSS/ptodata/1/pubpaa/US09_NEW_PUB.pep.*				
7: /SIDSS/ptodata/1/pubpaa/US09_NEW_PUB.pepl.*					8: /SIDSS/ptodata/1/pubpaa/US10_NEW_PUB.pep.*				
9: /SIDSS/ptodata/1/pubpaa/US10_NEW_PUB.pepl.*					10: /SIDSS/ptodata/1/pubpaa/US11_NEW_PUB.pep.*				
11: /SIDSS/ptodata/1/pubpaa/US11_NEW_PUB.pepl.*					12: /SIDSS/ptodata/1/pubpaa/US60_NEW_PUB.pep.*				
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.									
SUMMARIES									
Result No.	Score	Query Match	Length	ID	Description				
1	2673	96.2	507	11	US-11-033-039-96				
2	1502	54.0	481	9	US-10-508-263-28				
3	1499.5	54.0	485	9	US-10-508-263-22				
4	1496	53.8	495	9	US-10-508-263-20				
5	973.5	35.0	562	9	US-10-508-263-24				
					Sequence 96, Appl				
					Sequence 28, Appl				
					Sequence 22, Appl				
					Sequence 20, Appl				
					Sequence 24, Appl				

ALIGNMENTS									
RESULT 1									
US-11-033-039-96									
; Sequence 96, Application US/11033039									
; Publication No. US20060002947A1									
; GENERAL INFORMATION:									
; APPLICANT: HUMPHREYS, ROBERT									
; APPLICANT: XU MINZHEN									
; TITLE OF INVENTION: LI-KEY/ANTIGENIC EPIOTOPE HYBRID PEPTIDE VACCINES									
; FILE REFERENCE: REH-2017US01									
; CURRENT APPLICATION NUMBER: US/11/033,039									
; CURRENT FILING DATE: 2005-01-11									
; PRIOR APPLICATION NUMBER: 10/245,871									
; PRIOR FILING DATE: 2002-09-17									
; PRIOR APPLICATION NUMBER: 10/197,000									
; PRIOR FILING DATE: 2002-07-17									
; PRIOR APPLICATION NUMBER: 09/396,813									
; PRIOR FILING DATE: 1999-09-14									
; NUMBER OF SEQ ID NOS: 1452									
; SOFTWARE: PatentIn version 3.3									
; SEQ ID NO 96									
; LENGTH: 507									
; TYPE: PRT									
; ORGANISM: Arachis hypogaea									
US-11-033-039-96									
Alignment Scores:									
Pred. No.: 2.72e-220									
Score: 2673.00									
Percent Similarity: 100.0%									
Length: 507									
Matches: 507									
Conservative: 0									

Sequence 26, Appl
Sequence 62, Appl
Sequence 64, Appl
Sequence 30, Appl
Sequence 10, Appl
Sequence 18, Appl
Sequence 66, Appl
Sequence 12, Appl
Sequence 14, Appl
Sequence 68, Appl
Sequence 70, Appl
Sequence 60, Appl
Sequence 16, Appl
Sequence 22915, A
Sequence 22914, A
Sequence 22913, A
Sequence 25849, A
Sequence 25848, A
Sequence 17189, A
Sequence 25850, A
Sequence 12807, A
Sequence 12806, A
Sequence 89, Appl
Sequence 10, Appl
Sequence 907, Appl
Sequence 6597, Ap
Sequence 6596, Ap
Sequence 1818, Ap
Sequence 12808, A
Sequence 7436, Ap
Sequence 252, App
Sequence 249, App
Sequence 251, App
Sequence 248, App
Sequence 3385, Ap
Sequence 4181, Ap
Sequence 1, Appli
Sequence 387, App
Sequence 16091, A

Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 96.2% Indels: 0
 DB: 11 Gaps: 0

US-10-728-323-3 (1-1524) x US-11-033-039-96 (1-507)

QY	1	CGGCACCAACCGGAGAGAACGGTCCAGTTCAGCGCTCAATCGCGAGAGACTGAC	60
DB	1	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	20
QY	61	AATCGATTGAATCAGAGCGCGGTACATTGAGACTTGGAACCCCAACACAGGAGTTTC	120
DB	21	AsnArgLleGluSerGluGlyGlyTyrlleGluThrTrpAsnProAsnAsnGlnGluPhe	40
QY	121	GAATCGCGCGGCTGCCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCTT	180
DB	41	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro	60
QY	181	TTCTACTCCAATGCTCCCGAGGATCTTCATCCAGCAAGGAGGGGATACTTTGGTTG	240
DB	61	PheTySerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrlleGlyLeu	80
QY	241	ATATTCCCTGGTTGCTTAGACACTATGAAGAGCCTCACACAAGGTCTCGATCTCAG	300
DB	81	LlePheProGlyCysProArgHisTyrlleGluProHisThrGlnGlyArgSerGln	100
QY	301	TCCAAAGACCAAGAGCTTCTCAAGGAGAGACCAAGCCAAAGCCAAACAGAGATAGT	360
DB	101	SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnGlnArgAspSer	120
QY	361	CACCAGAGGTCCAGTTCGATGAGGTGATCTTCATTCAGTTCGCCCGGTGTGCT	420
DB	121	HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla	140
QY	421	TTCTGTCTCTACAACACACAGACACTGATGTGTGTGCTGTCTTCTTACTGACACCAAC	480
DB	141	PheTrpLeuTyraAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn	160
QY	481	AACAACGACAACAGCTTGATCAGTTCCTCCAGGAGATTCAAATTTGGCTGGGACACGGAG	540
DB	161	AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnThrGlu	180
QY	541	CAAGAGTTCTTAAGTTCAGCAGCAACAGCAGCAAGCAGCAGCAGAGAGCTTACCATAT	600
DB	181	GlnGluPheLeuArgTyrlleGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyrl	200
QY	601	AGCCCATACAGCCGCAAGTTCAGCTCAGCAAGAGAGCGTGAATTTAGCCCTCGAGGA	660
DB	201	SerProTySerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly	220
QY	661	CAGCAGCGCGCAGAGACGAGCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGATCTTC	720
DB	221	GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluGluGluGluGluGluGluGlu	240
QY	721	AGCGGCTTCACGCGGAGTTCCTGGAAACAGCCCTTCAGGTTGACACAGACAGATAGTG	780
DB	241	SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal	260
QY	781	CAAACTTAAGAGCGCAGACCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	840
DB	261	GlnAsnLeuArgGlyGluThrGluSerGluGluGluGluGluGluGluGluGluGluGlu	280
QY	841	GGCCTCAGATTTGAGCCGAGATAGAAAAGAGACGTGCCGACGAGAGAGAGAGAGAT	900
DB	281	GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrlleVal	300
QY	901	GAAGATCAATATCAATCAGATGAAGGATAGAGCGGTGGCAGGAGAGAGAGAGAGAG	960
DB	301	GluAspGluTyrlleGluTyrlleGluGluGluGluGluGluGluGluGluGluGluGlu	320
QY	961	GGGAATGGTATTCAAGAGACGATCTGCACCGCAAGTGTAAAAAGAACATTTGTAGAAC	1020
DB	321	GlyAsnGlyLleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn	340

QY	1021	AGATCCCCTGACATCTACAACCCCTCAAGCTGGTTCCTCACTCAAACTGCCAACGATCTCAAC	1080
DB	341	ArgSerProAspIleTyrlleTyrlleAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn	360
QY	1081	CTTCTAAATCTTAGGTGGCTTCGACCTAGTGTCTGAATATGGAAATCTCTACAGGAATGCA	1140
DB	361	LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrlleGlyAsnLeuTyrlleArgAsnAla	380
QY	1141	TTGTTTGTGCTCTACTACAACACCAACGACACAGCATCATATATCGATTGAGGGGACGG	1200
DB	381	LeuPheValAlaHisTyrlleAsnThrAsnAlaHisSerIleIleTyrlleArgLeuArgGlyArg	400
QY	1201	GCTCAGCTCAGTCTGTCGACAGCAGCAACGCGCAACAGAGTGTACGAGAGAGTCTCAAGG	1260
DB	401	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrlleAspGluGluLeuGlnGlu	420
QY	1261	GGTCAAGTCTGTGTCGCCAGAACTTCGCGCGTTCGCTGGAAAGTCCAGAGCGCAGAAC	1320
DB	421	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn	440
QY	1321	TTTCAATACGTGGCATTCAGACAGACTCAAGGCCAGCATAGCCAACTTCGCGCGGTGAA	1380
DB	441	PheGluTyrlleValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu	460
QY	1381	AACTCGCTCATAGATAACTTCGCGAGGAGGTGGTGCAAATTCATATGCGCTCCAAAGG	1440
DB	461	AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrlleGlyLeuGlnArg	480
QY	1441	GAGCAGGCAAGGCAGCTTAAGAACCAACACCCCTTCAAGTTCCTTCGCTCCACGCTCAG	1500
DB	481	GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProProSerGln	500
QY	1501	CAGTCTCCGAGGCTGTGGCT	1521
DB	501	GlnSerProArgAlaValAla	507

RESULT 2

US-10-508-263-28
 ; Sequence 28, Application US/10508263
 ; Publication No. US20050260754A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BASF Plant Science GmbH
 ; TITLE OF INVENTION: Constructs and methods for regulating gene expression

; FILE REFERENCE: 53262-20085.00
 ; CURRENT APPLICATION NUMBER: US/10/508,263
 ; CURRENT FILING DATE: 2004-09-20

; NUMBER OF SEQ ID NOS: 126
 ; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 28
 ; LENGTH: 481
 ; TYPE: PRT

; ORGANISM: Glycine max
 US-10-508-263-28

Alignment Scores:
 Pred. No.: 3,37e-120 Length: 481
 Score: 1502.00 Matches: 297
 Percent Similarity: 71.6% Conservative: 68
 Best Local Similarity: 58.2% Mismatches: 91
 Query Match: 54.0% Indels: 54
 DB: 7 Gaps: 7

US-10-728-323-3 (1-1524) x US-10-508-263-28 (1-481)

QY	1	CGGCACCAACCGGAGAGAACGGTCCAGTTCAGCGCTCAATCGCGAGAGACTGAC	60
DB	23	ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp	42
QY	61	AATCGATTGAATCAGAGCGCGGTTCATTTGAGACTTGGAAACCCCAACACAGGAGTTTC	120
DB	43	AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe	62

Db 120 SerGlnArgPro-----GlnAspArg 126
QY 361 CACAGAAAGGTGACCGTTTCGATGAGGGTGATCTCATTTGCGAGTTCCCAACCGGTGTGCT 420
Db 127 HisGlnLysValHisArgPheArgGluGlyAspLeuIleAlaValProThrGlyValAla 146
QY 421 TTCTGGCTCTAACAGCACACGACACTGATGTGTGTGCTGTTCTTCTTACTGACACCAAC 480
Db 147 TrpTrpMetTyrAsnAsnGluAspThrProValAlaValSerIleIleAspThrAsn 166
QY 481 AACACAGCAACACAGCTTGATCAGTTCCCGAGAGATTCAATTTGGCTGGGACACGGAG 540
Db 167 SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
QY 541 CAAGAGTTCTTAAGTTACAGCAACAAAGACAGACAAAGCAGACGAGAAAGCTTACCATAT 600
Db 187 GlnGluPheLeuLysTyrGlnGlnGln----- 195
QY 601 AGCCCATACAGCCCGCAAGTACGCTAGACAGAGAGCGGTGAATTTAGCCCTCGAGGA 660
Db 196 -----GlnGlnGlyGlySerGlnSerGlnLysGly 205
QY 661 CAGCACAGCCGACAGAACGACGACGACAGAAAGAAAGAAAGAAAGAGGTGGAACATCTTC 720
Db 206 LysGln-----GlnGluGluGluAsnGluGlySerAsnIleLeu 218
QY 721 AGCGGTTTCAGCCGAGTTCCTGGAAACAAGCCTTCAGGTTGACGACAGACAGATAGTG 780
Db 219 SerGlyPheAlaProGluPheLeuLysGluAlaPheGlyVal---AsnMetGlnIleVal 237
QY 781 CAAACCTTAAGGCGCAGACCGAGTGAAGAGGGGAGCATTTGTGACGTGAGGGGA 840
Db 238 ArgAsnLeuGlnGlyGluAsnGluGluAspSerGlyAlaIleValThrValLysGly 257
QY 841 GGCTCAGAAATCTGAGCCAGATAGAAAGAGACGTGCCGACGAAAGAGAAATACGAT 900
Db 258 GlyLeuArgValThrAlaProAlaMetArgLysProGlnGlnGluAspAspAsp 277
QY 901 GAAGATGAA-----TATGAATACGATGAAGAGGATAGAGGCGTGGCGGGA 948
Db 278 GluGluGluGlnProGlnCysValGluThrAspLysGlyCysGlnArgGlnSerLysArg 297
QY 949 AGCAGAGGAGGGGAATGGTATTGAAGACGATCTGCACCGCAAGTCTAAAAAGAAC 1008
Db 298 SerArg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgGlnAsn 314
QY 1009 ATTGGTAGAAACAGATCCCTGCATCTACACCTCAAGCTGGTTTCACTCAAAAATGCC 1068
Db 315 IleGlyGlnAsnSerSerProAspIleTyrAsnProGlnAlaGlySerIleThrAla 334
QY 1069 ACGATCTCAACCTTCTAATACATTAGGTGGCTTGGACCTAGTGTCTGAATATGGAATCTC 1128
Db 335 ThrSerLeuAspPheProAlaLeuTrpLeuLeuLysLeuSerAlaGlnTyrGlySerLeu 354
QY 1129 TACAGGAATGCAATGTTGTCGCTCACTCAACACCAACGACACAGCATCATATATCGA 1188
Db 355 ArgLysAlaAlaMetPheValProHisTyrThrLeuAsnAlaAsnSerIleIleTyrAla 374
QY 1189 TTCAGGGGACGGCTCACGTGCAAGTCGTGGACGACCAACGCGACAGAGTGTACGACGAG 1248
Db 375 LeuAsnGlyArgAlaLeuValGlnValValAsnCysAsnGlnGlyGluArgValPheAspGly 394
QY 1249 GAGCTTCAAGAGGTCACGTGCTTGTGGTCCCAAGAACTTCGCCGCTGCGTGGAAAGTCC 1308
Db 395 GluLeuGlnGluGlyValLeuIleValProGlnAsnPheAlaValAlaLaLysSer 414
QY 1309 CAGACCGAGACTTCCGAATACGTGGCATCAACACAGACTCAAGGCCCGCATAGCCAAC 1368
Db 415 GlnSerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsn 434
QY 1369 CTCGCCGTTGAAACCTCCGTATAGATAAACCCTCCGAGAGAGGTGGTTGCAATTCATAT 1428
Db 435 LeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnHisThrPhe 454

RESULT 4

US-10-508-263-20
; Sequence 20, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Glycine max
US-10-508-263-20

Alignment Scores:
Pred. No.: 1,116-119 Length: 495
Score: 1496.00 Matches: 292
Percent Similarity: 70.9% Conservative: 73
Best Local Similarity: 56.7% Mismatches: 100
Query Match: 53.8% Indels: 50
DB: Gaps: 6

US-10-728-323-3 (1-1524) x US-10-508-263-20 (1-495)

QY 1 CGGAGCAACCGGAGAGAACCGGTGCCAGTTCAGCGCCTCAATGCGCAGAGACTGAC 60
Db 23 ArgGluGlnProGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProAsp 42
QY 61 AATCGCATTAATCAGAGGGCGGTTCATATTGAGACTTGGAAACCCCAACACGAGGTTC 120
Db 43 AsnArgIleGluSerGluGlyGlyLeuIleGluThrTrpAsnProAsnAsnLysProPhe 62
QY 121 GAATGCGCGCGGTGCGCTCTCTCGTTAGTCTCTCCGCGCAACGCCCTTCGTAGGCT 180
Db 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 82
QY 181 TTCTACTCAATGCTCCCGAGAGATCTTCATCCAGCAAGGAGGATACTTTGGTTG 240
Db 83 SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyLysGlyIlePheGlyMet 102
QY 241 ATATTCCCTGGTTGTCTAGACACTATGAAGAGCCTCACACAAAGGTGTCGATCTCAG 300
Db 103 IleTyrProGlyCysProSerThrPheGluGluProGlnGlnProGlnArgGlyGln 122
QY 301 TCCAAAAGACCAACAGAGCGTCTCCAAGGAGAGCAACCAAGCCCAACGACGAGATAGT 360
Db 123 SerSerArgPro-----GlnAspArg 129
QY 361 CACCAAGGTGCACCGTTTCGATGAGGGTGATCTCATTTGAGTTCACCGGTGTGCT 420
Db 130 HisGlnLysIleTyrAsnPheArgGluGlyAspLeuIleAlaValProThrGlyValAla 149
QY 421 TTCTGGCTCTACAACGACACACGACACTGATGTTGTTGCTGTTCTTCTTACTGACACCAAC 480
Db 150 TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerIleIleAspThrAsn 169
QY 481 AACACGCAACACGAGCTTGATCAGTTCCCGAGAGATTCAATTTGGCTGGGAACACGGAG 540
Db 170 SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 189


```
Db 237 PheLeuAlaGlnSerPheAsnThrAsnGlu---AspIleAlaGluLysLeu-----Glu 253
QY ACGAGAGTGAAGAAGAGGAGCCATTGTGACAGTGAAGGGAGCGCTCAGAACTTTGAGC 858
   : : : : :
Db 254 SerProAspGluArgLysGlnIleValThrValGluGlyGlyLeuSerValIleSer 273
   : : : : :
QY 859 CCAGATAGAAGAGAGCTGCCGACGAAGAAGAGGAATACGATGAA----- 903
   : : : : :
Db 274 ProLysTrpGlnGlnGlnAspGluAspGluAspGluAspGluAspGlu 293
   : : : : :
QY 904 -----GATGAATAT 912
   : : : : :
Db 294 GlnIleProSerHisProProArgArgProSerHisGlyLysArgGluGlnAspGluAsp 313
   : : : : :
QY 913 GAATACGATGAAGAGATGAG---AGCGTGGCAGGGGAGAGAGGC----- 957
   : : : : :
Db 314 GluAspGluAspGluAspLysProArgProSerArgProSerGlnGlyLysArgAsnLys 333
   : : : : :
QY 957 ----- 957

Db 334 ThrGlyGlnAspGluAspGluAspGluAspGluAspGlnProArgLysSerArgGluTrp 353
QY 958 -----AGGGGG--- 963
   : : : : :
Db 354 ArgSerLysLysThrGlnProArgArgProArgGlnGluGluProArgGluArgGlyCys 373
   : : : : :
QY 964 -----AATGGTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAGAACATGTT 1014
   : : : : :
Db 374 GluThrArgAsnGlyValGluGluAsnIleCysThrLeuLysLeuHisGluAsnIleAla 393
   : : : : :
QY 1015 AGAAACAGATCCCTCACATCTACAACTTCAAGCTGTCTCAAGCTGTCTCAAACTGCCACGAT 1074
   : : : : :
Db 394 ArgProSerArgAlaAspPheTyrAsnProLysAlaGlyArgIleSerThrLeuAsnSer 413
   : : : : :
QY 1075 CTCAACTTCTAATACTTAGTGGCTTGGACCTAGTGTCTGAATATCTACAGG 1134
   : : : : :
Db 414 LeuThrLeuProAlaLeuArgGlnPheGlnLeuSerAlaGlnTyrValValLeuTyrLys 433
   : : : : :
QY 1135 AATGCATTTGTTCTCCTCATTACACACCAACGACAGCATCATATATCATGATTGAGG 1194
   : : : : :
Db 434 AsnGlyIleTyrSerProHisTrpAsnLeuAsnAlaAsnSerValIleTyrValThrArg 453
   : : : : :
QY 1195 GGACGGGCTCACGTCAAGTCGTGGACAGCAACGCGCAACAGAGTGTACGACGAGGAGCTT 1254
   : : : : :
Db 454 GlyGlnGlyLysValArgValValAsnCysGlnGlyAsnAlaValPheAspGlyGluLeu 473
   : : : : :
QY 1255 CAAGAGGGTCACGTGCTTGTGTGCCACAGAACTTCGCGCTCGCTGGAAGTCCACAGAGC 1314
   : : : : :
Db 474 ArgArgGlyGlnLeuLeuValValProGlnAsnPheValValAlaGluGlnAlaGlyGlu 493
   : : : : :
QY 1315 GAGAACTTGAATACGTGGCATTCAGACAGACTCAAGGCCAGCATAGCCACCTCGCC 1374
   : : : : :
Db 494 GlnGlyPheGluTyrIleValPheLysThrHisAsnAlaValThrSerTyrLeu--- 512
   : : : : :
QY 1375 GGTGAAACTCCCTCATAGATAACCTCCGGAGAGGTGGTGTGCAATTCATATGSCCTC 1434
   : : : : :
Db 513 -----LysAspValPheArgAlaIleProSerGluValLeuAlaHisSerTyrAsnLeu 530
   : : : : :
QY 1435 CAAAGGAGCAGCAAGGAGCTTAAGAAACAACACCCCTTCAAGTTCCTCGTT---CCA 1491
   : : : : :
Db 531 ArgGlnSerGlnValSerGluLeuLysTyrGluGlyAsnTrpGlyProLeuValAsnPro 550
   : : : : :
QY 1492 CCGTCTCAGAG---TCTCCGAGG 1512
   : : : : :
Db 551 GluSerGlnGlnGlySerProArg 558
```

RESULT 6

```
US-10-508-263-26
; Sequence 26, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
```

```
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 516
; TYPE: PR1
; ORGANISM: Glycine max
US-10-508-263-26

Alignment Scores:
Pred. No.: 5,42e-71 Length: 516
Score: 926.50 Matches: 207
Percent Similarity: 53.1% Conservative: 72
Best Local Similarity: 39.4% Mismatches: 163
Query Match: 33.3% Indels: 83
DB: 9 Gaps: 11
```

```
US-10-728-323-3 (1-1524) x US-10-508-263-26 (1-516)

QY 19 AACGGCTGCCAGTTCAGCGCCTCAATGCGCAGAGACCTGACAATGCGCATTTGAATCAGAG 78
   : : : : :
Db 31 AsnGluCysGlnLeuAsnAsnLeuAsnAlaLeuGluProAspHisArgValGluSerGlu 50
   : : : : :
QY 79 GCGGGTTACATTGAGACTTGGAAACCCCAACACCAAGGAGTTCGAATGCCCGCGCTCGCC 138
   : : : : :
Db 51 GlyGlyLeuIleGluThrTrpAsnSerGlnHisProGluLeuGlnCysAlaGlyValThr 70
   : : : : :
QY 139 CTCCTCGGTTAGTCTCTCGCGCAACGCGCTTCGTAGGCTTCTTACTCTCAATGCTCCC 198
   : : : : :
Db 71 ValSerLysArgThrLeuAsnArgAsnGlySerHisLeuProSerTyrLeuProTyrPro 90
   : : : : :
QY 199 CAGGAGATCTTCATCCAGCAAGAGGGGATACTTTGGGTTGATATTCCTGGTTGCTCT 258
   : : : : :
Db 91 GlnMetIleIleValValGlnGlyLysGlyAlaIleGlyPheAlaPheProGlyCysPro 110
   : : : : :
QY 259 AGACACTATGAAGAGCCTCACACACAGGTCTCGATCTCAGTCCCAAGACCAACCAAGA 318
   : : : : :
Db 111 GluThrPheGluLysProGlnGlnGlnSerSerArg----- 122
   : : : : :
QY 319 CGTCTCCAAGAGAGAAGACCAAGAGCAACAGACAGATAGTACCAGAAAGGTGACCGT 378
   : : : : :
Db 123 -----ArgGlySerArgSerGlnGlnGlnLeuGlnAspSerHisGlnLysIleArgHis 140
   : : : : :
QY 379 TTCGATAGGGTGATCTCATTCAGCTTCCACCGGTGTTGCTTTCTGGCTCTTACAAACGAC 438
   : : : : :
Db 141 PheAsnGluGlyAspValLeuValIleProLeuGlyValProTyrTrpThrTyrAsnThr 160
   : : : : :
QY 439 CAGGACACTGATGTTGCTCTTCTTCTTACTGACACCAACCAACCAACCAACCAACGCTT 498
   : : : : :
Db 161 GlyAspGluProValValAlaIleSerProLeuAspThrSerAsnPheAsnGlnLeu 180
   : : : : :
QY 499 GATCAGTTCCTCCAGGAGATTCAATTTGGCTGGGAACACGAGCAAGAGTTCCTTAAGGTAC 558
   : : : : :
Db 181 AspGlnAsnProArgValPheTyrLeuAlaGlyAsnProAspIleGlu----- 196
   : : : : :
QY 559 CAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATATAGCCCATACAGCCCGCAA 618
   : : : : :
Db 197 -----HisProGluThrMetGln 202
   : : : : :
QY 619 AGTCAGCCTTAGACAAGAGCGTGAATTTAGCCCTCGAGGACAGCAGACCGCCAGAGAA 678
   : : : : :
Db 203 GlnGlnGlnGlnGlnLysSerHisGlyGlyArgLysGlnGlyGlnHis----- 218
   : : : : :
QY 679 CGAGCAGGACAAGAAGAAAGAAACATCTTCAGCGGGTTCACGCGGGAG 738
   : : : : :
Db 219 -----ArgGlnGlnGlnGluGluGlyGlySerValLeuSerGlyPheSerLysHis 235
   : : : : :
QY 739 TTCCTGGAACAGCCCTTCAGGTTGACGACACAGATAGTGCAAAACCTTAAGACGGCAG 798
   : : : : :
Db 236 PheLeuAlaGlnSerPheAsnThrAsnGlu---AspThrAlaGluLysLeuArg----- 252
   : : : : :
```


Qy	799	ACCAGAGTGAAGAAGAGGGAGCCATTGTGACAGTGAAGGAGGAGCCCTCAGAAATCTTGAGC	858
Db	253	SerProAspAspGluArgLysGlnIleValThrValGluGlyLeuSerValIleSer	272
Qy	859	CCACATAGAAGACAGCAGCTGCCGACGACGAAGAGAGAAATACGATGAA	903
Db	273	ProLysTrpGlnGluGlnGluAspGluAspGluAspGluGluTrpGlyArgThr	292
Qy	904	-----GATGAATATGAATAC	918
Db	293	ProSerTyrProProArgArgProSerHisGlyLysHisGluAspAspGluAspGluAsp	312
Qy	919	GATCAAGAGGATAGA	933
Db	313	GluGluGluAspGlnProArgProAspHisProProGlnArgProSerArgProGluGln	332
Qy	934	-----AGGCGTGCAGGGGAAGCAGAGGAGGGGGAATGATTGAAGAGAGCATCTGC	987
Db	333	GlnGluProArgGlyArgGlyCysGlnThrArg--AsnGlyValGluGluAsnIleCys	351
Qy	988	ACCSCAAGTGTAAAAGACATTGGTGTAGAAACAGATCCCTGACATCTCAACCCCTCAA	1047
Db	352	ThrMetLysLeuHisGluAsnIleAlaArgProSerArgAlaAspPheTyrAsnProLys	371
Qy	1048	GCTGGTTCACTCAAAACTGCCCAACGATCTCAACCTTCTAATACTAGCTGGCTTGGACCT	1107
Db	372	AlaGlyIleSerThrLeuAsnSerLeuThrLeuProAlaLeuArgGlnPheGlyLeu	391
Qy	1108	AGTGCTGAATATGAAATCTCTACAGGAATGCATTTGTTGCGTCTACTACACACCAAC	1167
Db	392	SerAlaGlnTyrValValLeuTyrArgAsnGlyIleTyrSerProAspTrpAsnLeuAsn	411
Qy	1168	GCACACAGCATCATATATCGATTGAGGGGACGGGCTCACGTGCAAGTCGTGGACAGCAAC	1227
Db	412	AlaAsnSerValThr---MetThrArgGlyLysGlyArgValArgValValAsnCysGln	430
Qy	1228	GGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGCTCACGTGCTGTGTGGTCCACAGAAC	1287
Db	431	GlyAsnAlaValPheAspGlyGluLeuArgGlyGlnLeuLeuValValProGlnAsn	450
Qy	1288	TTCCGCTCGCTGGAAGTCCCGAGGCGAGAACCTTCGAATACGTGGCATCAAGACAGAC	1347
Db	451	ProAlaValAlaGluGlnGlyGlyGluGlnGlyLeuGluTyrValValPheLysThrHis	470
Qy	1348	TCAGGGCCACGATAGCCACCTCGCCGGTGAAACTCCGTCATAGATAACCTGCCGGAG	1407
Db	471	HisAsnAlaValSerSerTyrIle-----LysAspValPheArgValIleProSer	487
Qy	1408	GAGTGGTTGCAAAATTCATATGGCTCTCAAAGGAGGAGGCAAGCGCAGCTTAAG-----	1461
Db	488	GluValLeuSerAsnSerTyrAsnLeuGlyGlnSerGlnValArgGlnLeuLysTyrGln	507
Qy	1462	---AACAAACACCCC	1473
Db	508	GlvAsnSerGlyPro	512

RESULT 7

```

RESUBMIT /
US-10-508-263-62
; Sequence 62, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 62
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-508-263-62

```

Alignment Scores:	1.93e-58	Length:	472
Pred. No.:	779.50	Matches:	176
Percent Similarity:	51.6%	Conservative:	88
Best Local Similarity:	34.4%	Mismatches:	175
Query Match:	28.0%	Indels:	73
DB:	9	Gaps:	10
US-10-728-323-3 (1-1524) x US-10-508-263-62 (1-472)			
QY	4	CAGCAACCGGAGAG-----AACCGCTGCCAGTTCACGAGCCCTCAATGCCAGAGACCT	57
DB	27	GlnGlnGlnGlnPheProAsnGluCysGlnLeuAspGlnLeuAsnAlaLeuGluPro	46
QY	58	GACAAATCCGATTTGAATCAGAGGGCGGTTACATTCAGACTTGGAAACCCCAACACAGAGG	117
DB	47	SerHisValLeuLeuSerGluAlaGlyArgIleGluValTrpAspHisHisAlaProGln	66
QY	118	TTCAATTCGCCGGCGCTCTCGCTTAGTCTCCGCCCAACGCCCTTCGTAGG	177
DB	67	LeuArgCysSerGlyValSerPheAlaArgTyrIleIleGluSerLysGlyLeuTrpLeu	86
QY	178	CCTTTCTACTCCAATGCTCCCGAGAGATCTTCATCCAGCAAGAAAGGGGATCTTTGGG	237
DB	87	ProSerPhePheAsnThrAlaLysLeuSerPheValAlaLysGlyArgGlyLeuMetGly	106
QY	238	TTGATATTCCTGGTTGCTCTAGACACTATGAAGAGCGCTCACACAAGGTCTGTCGATCT	297
DB	107	LysValIleProGlyCysAlaGluThrPheGlnAsp-----	118
QY	298	CAGTCCCAAGACCACCAAGACGCTCCAGGAGACCAACCAACACACAGAGAT	357
DB	119	SerSerGluPheGlnPro---ArgPheGluGlyGlnGlyGlnSerGlnArgPheArgAsp	137
QY	358	AGTCACAGAGGTGCACCGTTTCGATGAGGGTGATCTTCATTCGAGTTCCCAACCGGTGT	417
DB	138	MethIleGlnLysValGluHisIleArgSerGlyAspThrIleAlaThrThrProGlyVal	157
QY	418	GCTTTCTGGCTTCAACAGCACGACACACTGATGTTGTGTGTTCTTCTTACTGACACC	477
DB	158	AlaGlnTrpPheTyrAsnAspGlyGlnProLeuValIleValSerValPheAspLeu	177
QY	478	AACAAACAGCAACACGCTTGATCAGTTCCTCCAGGAGATTCAATTCGGCTGGGAACAG	537
DB	178	AlaSerHisGlnAsnGlnLeuAspArgAsnProArgProPheTyrLeuAlaGlyAsn---	196
QY	538	GAGCAAGAGTTCTTAAGGTACCAACAACAGACAGCAAGACGACGAAGAGCTTACCA	597
DB	196	-----	196
QY	598	TATAGCCCATACCGCGCAAGTACGCCTAGACAAGAGAGCGTGAATTTAGCCCTCGA	657
DB	197	-----AsnProGlnGlyGlnValTrpLeuGlnGlyArgGluGlnGlnProGln	212
QY	658	GGACAGCACAGCCGCAGAGAACGACGAGGACAAGAAAGAAACGAAGGTGGAAACATC	717
DB	213	-----LysAsnIle	215
QY	718	TTCAGCGGCTTACCGCGAGTTCTCTGAAACAAGCCTTCCAGTTGACGACGACAGATA	777
DB	216	PheAsnGlyPheGlyProGluValIleAlaGlnAlaLeuLysIle---AspLeuGlnThr	234
QY	778	GTGCAAACCTTAAGAGGCGACCGAGAGTCAAGAGGGAGCGCATTTGTACAGTGAGG	837
DB	235	AlaGlnGlnLeu-----GlnAsnGlnAspAsnArgGlyAsnIleValArgValGln	252
QY	838	GGAGCGCTCACAATCTTGAGCGCCAGATAGAAAGAGACGTCGCCGACGAAGAGAAATAC	897
DB	253	GlyProPheGlyValIleArgProProLeuArgGlyGlnArgProGlnGluGlu---	271
QY	898	GATGAAGATGAATATGAATACGATGACGAGTATGAGCGCTGCAGCGGGAACGACGAGC	957

Db 272 -----GluGluGluGlyArgHisGlyArg----- 279
 QY 958 AGGGGAATGATTGAAGAGACGATCTGCACCGCAAGTGTAAAGAACATCTGTGTAGA 1017
 Db 280 HisGlyAsnGlyLeuGluGluThrIleCysSerAlaArgCysThrAspAsnLeuAspAsp 299
 QY 1018 AACAGATCCCTGACATCTACAACTCAAGCTGTCTACTCAAAATGCTCCACGATCTC 1077
 Db 300 ProSerArgAlaAspValTyrIysProGlnLeuGlyTyrIleSerThrLeuAsnSerTyr 319
 QY 1078 AACCTTCTAATCTAGGTGGCTTGGACCTAGTGTGTAATATGAAATCTCTACAGAAAT 1137
 Db 320 AspLeuProIleLeuArgPheIleArgLeuSerAlaLeuArgGlySerIleArgGlnAsn 339
 QY 1138 GCAATTTTGTGCTCACTACACACCAACGACACGATCATATATCGATTGAGGGA 1197
 Db 340 AlaMetValLeuProGlnTrpAsnAlaAsnAlaIleLeuTyrGluThrAspGly 359
 QY 1198 CGGGCTCAGTGCAGTCTGGAGACAGCAACGGCAACAGAGTGTAACGACGAGGAGCTTCAA 1257
 Db 360 GluAlaGlnIleGlnIleValAsnAspAsnGlyAsnArgValPheAspGlyGlnValSer 379
 QY 1258 GAGGGTCACGTCTGTGTGGCCACAGAACTTCGCGTGTGGTGAAGTCCACAGACGAG 1317
 Db 380 GlnGlyGlnLeuIleAlaValProGlnGlyPheSerValValIysArgAlaThrSerAsn 399
 QY 1318 AACTTCGAATAGTGCATTCAGACAGACTCAAGGCCCAGACATAGCCAACTCGCGGT 1377
 Db 400 ArgPheGlnTrpValGluPheIysThrAsnAlaAsnAlaGlnIleAsnThrLeuAlaGly 419
 QY 1378 GAAAACTCCGTCTATAGTAACCTGCGGAGGAGTGTGGCAAAATTCATATGSCCTCCAA 1437
 Db 420 ArgThrSerValLeuArgGlyLeuProLeuGluValIleThrAsnGlyPheGlnIleSer 439
 QY 1438 AGGGAGCAGCAGGAGCTTAAGAACAAACCCCTTCAAGTCTCTCGTT----- 1488
 Db 440 ProGluAlaArgArgValIysPheAsnThrLeuGluThrThrLeuThrHisSerSer 459
 QY 1489 ---CCACGCTCTCAGCAGTCTCCGAGGGCTGTGGCT 1521
 Db 460 GlyProAlaSerTyrGlyArgProArgValAlaAla 471

RESULT 8

US-10-508-263-64
 ; Sequence 64, Application US/10508263
 ; Publication No. US20050260754A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BASF Plant Science GmbH
 ; TITLE OF INVENTION: Constructs and methods for regulating gene expression
 ; FILE REFERENCE: 53262-20085.00
 ; CURRENT APPLICATION NUMBER: US/10/508,263
 ; CURRENT FILING DATE: 2004-09-20
 ; NUMBER OF SEQ ID NOS: 126
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 64
 ; LENGTH: 472
 ; TYPE: PR1
 ; ORGANISM: Arabidopsis thaliana
 US-10-508-263-64

Alignment Scores:

Pred. No.:	3,49e-58	Length:	472
Score:	776.50	Matches:	176
Percent Similarity:	51.6%	Conservative:	88
Best Local Similarity:	34.4%	Mismatches:	175
Query Match:	27.9%	Indels:	73
DB:	9	Gaps:	10

US-10-728-323-3 (1-1524) x US-10-508-263-64 (1-472)

QY 4 CAGCAACCGGAGGAG-----AACGGGTGCCAGTCTCCAGCGCTCAATGCGCAGACCT 57
 Db 27 GlnGlnGlyGlnGlnPheProAsnGluCysGlnLeuAspGlnLeuAsnAlaLeuGluPro 46

QY 58 GACAATCGCATTTGAATCAGAGGCGGTTTACATTGAGACTTGGAAACCCCAACCAACGAGGAG 117
 Db 47 SerHisValLeuLeuSerGluAlaGlyArgIleGluValTrpAspShiHisAlaProGln 66
 QY 118 TTTCGAATGCGCGCGGTGCGCTCTCTCGCTTAGTCTCCGCCGCAACCCCTTCGTAGG 177
 Db 67 LeuArgCysSerGlyValSerPheAlaArgTyrIleIleGluSerIysGlyLeuTyrLeu 86
 QY 178 CTTTCTTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGGGATCTTTGGG 237
 Db 87 ProSerPhePheAsnThrAlaIysLeuSerPheValAlaIaLysGlyArgGlyLeuMetGly 106
 QY 238 TTGATATTCCTGGTGTGCTAGACACTATATGAAGAGCCTCACACACAAGTCGTGATCT 297
 Db 107 LysValIleProGlyCysAlaGluThrPheGlnAsp----- 118
 QY 298 CAGTCCCAAGACCCACAGACGCTCTCCAAGGAGAGACCAAGCAACAGCAACAGAGAT 357
 Db 119 SerSerGluPheGlnPro---ArgPheGluGlyGlnGlyGlnSerGlnArgPheArgAsp 137
 QY 358 AGTCACCAAGAGTGCACCGTTTCGATGAGGCTGATCTCATTCGAGTTCCTCCACCGTGT 417
 Db 138 MetHisGlnLysValGluHisIleArgSerGlyAspThrIleAlaThrThrProGlyVal 157
 QY 418 GCTTTTCTGCTCTACAACGACCACGACACTGATGTTGTTGCTGTTTCTTCTTACTGACACC 477
 Db 158 AlaGlnTrpPheTyrAsnAspGlyGlnGluProLeuValIleValSerValPheAspLeu 177
 QY 478 AACAAACAACGACACAGCAGCTTCAGTTCCTCCAGGAGATTCAAATTTGGCTGGGAACAG 537
 Db 178 AlaSerHisGlnAsnGlnLeuAspArgAsnProArgProPheTyrLeuAlaGlyAsn--- 196
 QY 538 GAGCAAGAGTCTTAAAGTTACCAGCAACAAAGCAGACAAAGCAGACGAGAAAGCTTACCA 597
 Db 196 ----- 196
 QY 598 TATAGCCCATACAGCCCGCAAGTCAGCCTAGACAAAGAGCGTGAATTTAGCCCTCGA 657
 Db 197 -----AsnProGlnGlyGlnValTrpLeuGlnGlyArgGlnGlnProGln 212
 QY 658 GGACAGCACAGCGCGCAGAGACGAGCAGGACAAAGAAAGAAACGAGGTGAAACATC 717
 Db 213 -----LysAsnIle 215
 QY 718 TTCAGCGGCTTCACGCGGAGTTCCTGGAAACAGCCTTCAGGTTGACGACGACAGATA 777
 Db 216 PheAsnGlyPheGlyProGluValIleAlaGlnAlaLeuIysIle---AspLeuGlnThr 234
 QY 778 GTGCAAAACCTAAGAGCGCAGACCGAGAGTGAAGAGAGGAGCGCCATTCTGCAGGTGAGG 837
 Db 235 AlaGlnGlnLeu-----GlnAsnGlnAspAsnArgIlyAsnIleValArgValGln 252
 QY 838 GGAGGCTTCAGAAATCTTGAGCCCATAGAAAAGACGTCGCCGACGAAGAAGAGGAAATAC 897
 Db 253 GlyProPheGlyValIleArgProProLeuArgGlyGlnArgProGlnGluGlu--- 271
 QY 898 GATGAAGATGAATATGAATACGATGAAGAGGATAGAGCGGTGGGGGAGGAGCAGAGGC 957
 Db 272 -----GluGluGluGlyArgHisGlyArg----- 279
 QY 958 AGGGGAATGTTTGAAGACGATCTGCACCGCAAGTGTAAAGAAACATTTGTTGTA 1017
 Db 280 HisGlyAsnGlyLeuGluGluThrIleCysSerAlaArgCysThrAspAsnLeuAspAsp 299
 QY 1018 AACAGATCCCTGACATCTACAACTCAAGCTGTGTTCTACTCAAACTGCTCCACGATCTC 1077
 Db 300 ProSerArgAlaAspValTyrIysProGlnLeuGlyTyrIleSerThrLeuAsnSerTyr 319
 QY 1078 AACCTTCTAATCTAGGTGGCTTGGACCTAGTGTGAATATGGAATCTCTACAGAAAT 1137
 Db 320 AspLeuProIleLeuArgPheIleArgLeuSerAlaLeuArgGlySerIleArgGlnAsn 339

```

QY 1138 GCATTGTTGTCGCTCACTACACACCAACGACACGATCATATATCGATTGAGGGA 1197
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
340 AlaMetValLeuProGlnTrpAsnAlaAsnAlaLeuThrValThrAspGly 359
QY 1198 CGGGCTCAGTCGAGTGTGACAGCAACGCGCAACAGAGTGTACGAGGAGCTTCAA 1257
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
360 GluAlaGlnIleGlnIleValAsnAspAsnGlyAsnArgValPheAspGlyValSer 379
QY 1258 GAGGTCACGTCGTTGTGTGCGCACAGAACTTCGCGCTGCGTGGAAAGTCCCGAGCGAG 1317
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
380 GlnGlyGlnLeuIleAlaValProGlnGlyPheSerValValYsArgAlaThrSerAsn 399
QY 1318 AACTTCGAATACGTGCATTCAGACAGACAGTCAAGGCCCGAGCATAGCCCAACCTCCCGGT 1377
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
400 ArgPheGlnTrpValGluPheLysThrAsnAlaAsnAlaGlnIleAsnThrLeuAlaGly 419
QY 1378 GAAACTCGCTCATAGATAACCTCGGAGGAGGTGTGCAAAATTCATATGCGCTCCAA 1437
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
420 ArgThrSerValLeuArgGlyLeuProLeuGluValIleThrAsnGlyPheGlnIleSer 439
QY 1438 AGGAGCAGCGCAAGCGAGCTTAAGAACCAACACCCCTTCAAGTTCCTCGTT----- 1488
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
440 ProGluGluAlaArgValIleYsPheAsnThrLeuGluThrLeuThrHisSerSer 459
QY 1489 ---CCACCGTCTCAGCAGTCTCCGAGGCGTGTGGCT 1521
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
460 GlyProAlaSerTyrglyArgProArgValAlaAla 471

RESULT 9
US-10-508-263-30
; Sequence 30, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; TYPE: PRN
; ORGANISM: Helianthus annuus
US-10-508-263-30

Alignment Scores:
Pred. No.: 7,06e-58 Length: 493
Score: 773.00 Matches: 186
Percent Similarity: 53.1% Conservative: 79
Best Local Similarity: 37.3% Mismatches: 170
Query Match: 27.8% Indels: 64
DB: 9 Gaps: 12

US-10-728-323-3 (1-1524) x US-10-508-263-30 (1-493)
QY 1 CGCGAGCAACCGAGGAGACCGTCCAGTTCGAGCGCTCAATGCGCAGACCTGAC 60
Db 24 GlnArgGlnGlnGlnAsnGlnCysGlnLeuGlnAsnIleGluAlaLeuProIle 43
QY 61 AATCGCATTAATTCAGAGCGCGGTGTACATTGAGACTTGGAAACCCCAACACGAGGATTC 120
Db 44 GluValIleGlnAlaGluAlaGlyValThrGluIleTrpAlaThrAspGlnPhe 63
QY 121 GAATGCGCGCGC-----GTGCGCTCTCTCGCTTA 150
Db 64 GlnCysAlaTrpSerIleLeuPheAspThrGlyPheAsnLeuValAlaPheSerCysLeu 83
QY 151 GTCCCTCGCGCAACGCCCTTCGTAGGCGCTTCTACTCAATGCTCCCGAGGAGATCTTC 210
Db 84 -----ProThrSerThrProLeuPheTrpProSerSerArgGlu----- 96
QY 211 ATCCAGCAAGGAAGGGGACTTTGGTGTGATATTCCCTGTTGTCCTAGACACTATGAA 270

```

```

Db 97 -----GlyValIleLeuProGlyCysArgThrTyrglu 108
QY 271 GAGCTCACACACAGGTGCTCGATCTCAGTCCCAAGACACCAAGACGCTCTCCAAAGA 330
Db 109 TyrSerGlnGluGln-----GlnPheSerGlyGluGlyGlyArgArgGlyGly 125
QY 331 GAAGACCAAAACCAACAGCAACAGATAGTCACCAAGAGGTGCACCGTTTCGATGAGGT 390
Db 126 GluGlyThrPhe-----ArgThrValIleArgLysLeuGluAsnLeuLysGluGly 142
QY 391 GATCTCATGTCAGTTCACCGGTGCTTCTTCTGCTCTTACCAAGCAGCACGACACATGAT 450
Db 143 AspValValAlaIleProThrGlyThrAlaHisTrpLeuHisAsnAspGlyAsnThrGlu 162
QY 451 GTTGTGCTGTTCTTCTTACTGACACCAACCAACAGCAACAGCTTGATCAGTTCCTCCC 510
Db 163 LeuValValValPheLeu---AspThrGlnAsnHisGluAsnGlnLeuAspGlnGln 181
QY 511 AGGAGATTCAATTTGGTGGAAACACGAGCAAGAGTTCCTTAAGGTACCAGCAACAAAGC 570
Db 182 ArgArgPhePheLeuAlaGlyAsnProGlnAlaGlnAlaGlnSerGlnGlnGlnGln 201
QY 571 AGACAAAGCAGCAAGAAAGCTTACCATATAGCCCATACAGCCCAAGAGCTCAGCTAGA 630
Db 202 ArgGlnProArgGln-----GlnSerProGlnArg 211
QY 631 CAAGAAGAGCGTGAATTTAGCCCTCGAGGACAGCACAGCCGCGAGAGACGAGACAA 690
Db 212 GlnArgGlnArgGlnArgGlnGlnGlnGlnGlnAla----- 224
QY 691 GAAGAAGAAAAACGAGGTGGAACATCTTCAGCGCTTCAGCGCGGAGTTCCTGGAACAA 750
Db 225 -----GlyAsnIlePheAsnGlyPheThrProGluLeuIleAlaGln 238
QY 751 GCCTTCAGTTCAGCAGACAGATAGTCAAAACCTTAAGAGGCGAGCCGAGAGTGAA 810
Db 239 SerPheAsnVal---AspGlnGluThrAlaGlnLysLeuGlnGly-----GlnAsnAsp 255
QY 811 GAAGAGGAGCGCATTTGTCACAGTGGGGGCGCTCAGAATCTTCAGCCCGAGATGAAAG 870
Db 256 GlnArgGlyHisIleValAsnValGlyGlnAspLeuGlnIleValArgProProGlnAsp 275
QY 871 AGACGTGCC-----GACAAAGAGAGAAATACGATGAAGATGAATATGAATACGATGAA 924
Db 276 ArgArgSerProArgGlnGlnGlnGlnAlaThrSerProArgGlnGlnGlnGln 295
QY 925 GAGGATAGAGCGGTGCGAGGGAGAGCAGAGCGGGGGAATGTTATTGAAGAGACGATC 984
Db 296 GlnGlnGlyArgArgGly-----GlyTrpSerAsnGlyValGluGluThrIle 311
QY 985 TGCACCGCAAGTGTATAAAGAACATTTGGTAGAACACAGATCCCTCGACATCTACCAACCT 1044
Db 312 CysSerMetLysPheLysValAsnIleAspAsnProSerGlnAlaAspPheValAsnPro 331
QY 1045 CAAGCTGTTCTCACTCAAAACCTGCCAACGATCTCAACCTTCTAATATCTTAGTGGCTTGA 1104
Db 332 GlnAlaGlySerIleAlaAsnLeuAsnSerPheLysPheProIleLeuGluHisLeuArg 351
QY 1105 CCTAGTGTCTGAATATGGAATCTCTACAGGAATGCAATGTTGTCGCTCCTACACACACC 1164
Db 352 LeuSerValGluArgGlyGluLeuArgProAsnAlaIleGlnSerProHisTrpThrIle 371
QY 1165 AACGCACACGATCATATATCGATTGAGGGAGCGGCTCAGCTCAAGTCTGTGACAGC 1224
Db 372 AsnAlaHisAsnLeuLeuTyrrValThrGluGlyAlaLeuArgValGlnIleValAspAsn 391
QY 1225 AACGCCAACAGAGTGTACGAGGAGCTTCAAGAGGCTTCAGCTGCTTGTGTGTCACAG 1284
Db 392 GlnGlyAsnSerValPheAspAsnGluLeuArgGluGlyGlnValValIleProGln 411
QY 1285 AACTTCGCGCTGCTGGAAAGTCCAGAGCGAGAACTTCGAATACGTCGATTCAGACA 1344
Db      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

```



```
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Brassica napus
US-10-508-263-18

Alignment Scores:
Pred. No.: 5,21e-55 Length: 509
Score: 739.50 Matches: 164
Percent Similarity: 49.9% Conservative: 91
Best Local Similarity: 32.1% Mismatches: 169
Query Match: 26.6% Indels: 87
DB: 9 Gaps: 8

US-10-728-323-3 (1-1524) x US-10-508-263-18 (1-509)
QY 19 AACGGTGCAGTTCAGCGCTCAATGCGCAGAGACCTGACATCGCATTTGAATCAGAG 78
DB 35 AsnAlaCysAsnLeuAspValLeuGlnProThrGluThrIleLysSerGlu 54
QY 79 GCGGTTACATTCAGACTTGGAAACCCCAACACACAGGAGTTTCGATGCGCGCTGCC 138
DB 55 AlaGlyArgValGluTyrTrpAspHisAsnAsnProGlnIleArgCysAlaGlyValSer 74
QY 139 CTCTCTCGTTCAGTCTCCGCGCAACGCGCTTCGTAGGCTTTTCTACTCCAAATGCTCC 198
DB 75 ValSerArgValIleIleGlnGlyGlyLeuTyrLeuProThrPheSerSerPro 94
QY 199 CAGGAGATCTTCATCAGCAAGAAAGGGGATACTTTGGGTTCATATCCCTGGTTGTCT 258
DB 95 LysIleSer-Tyr-ValValGlnGlyMetGlyIleSerGlyArgValValProGlyCysAla 114
QY 259 AGACACTAT-----GAAGAGCT-----CACACACAGGTGCTGCATCTCAGTCC 303
DB 115 GluThrPheMetAspSerGlnProMetGlnGlyGlnGlnGlyGlnProTrpGlnGly 134
QY 304 CAAGACCAACCAAGAGCTCTCAAGGAGAA-----GACCAAGCCAAACAGCAA-----CGA 354
DB 135 GlnGlnGlyGlnGlnGlyGlnGlnGlyGlnGlnGlyGlnGlnGlyGlnGlnGln 154
QY 334 -----GACCAAGCCAAACAGCAA-----CGA 354
DB 155 GlyGlnGlnGlyGlnGlnGlyGlnGlnGlyGlnGlnGlyGlnGlnGlyPheArg 174
QY 355 GATAGTCACCAAGGTGCACCGTTTCGATGAGGTGATCTCATTTGCAGTCCACCGGT 414
DB 175 AspMetHisGlnLysValGluHisValArgHisGlyAspIleAlaIleThrAlaGly 194
QY 415 GTTGCTTCTGCTCTACACACACACACACACTGATGTTGTTGTTCTTCTTACTGAC 474
DB 195 SerSerHisTrpIleTyrAsnThrGlyAspGlnProLeuValIleIleCysLeuLeuAsp 214
QY 475 ACCAACAACAACACACACAGCTTGATCTCCCGAGGAGATTCATATTTGGCTGGGAAC 534
DB 215 IleAlaAsnTyr-GlnAsnGlnLeuAspArgAsnProArgThrPheArgLeuAlaGlyAsn 234
QY 535 ACGGAGCAGAGTTCTTAGGTACCGCAACCAACAGCACAAGCAGACGAGAAGACTTGA 594
DB 235 AsnProGlnGlyGlySerGlnGlnGlnGlnGlnGlnGln----- 248
QY 595 CCATATAGCCCATACAGCCCGCAAAAGTCAGCTAGACAGAAGAGCGTGAATTTAGCCCT 654
DB 248 ----- 248
QY 655 CGAGGACACACACAGCGCCAGAGACGAGCAGGACAGAGAGAGAGAGAGAGAGAGAGAG 714
```

```
Db 249 -----Asn 249
QY 715 ATCTTCAGCGCTTCACGCGGAGTTCCTGGAAACAAGCTTCCAGGTTGACGACACAG 774
DB 250 MetLeuSerGlyPheAspProGlnValLeuAlaLeuLysIleAspValArg--- 268
QY 775 ATAGTCAAAAACCTAAGAGCGCAGACCGAGAGTGAAGAAGAGGAGCCATTGTGACAGTG 834
DB 269 LeuAlaGlnGluLeuGlnAsnGlnAspSer-----ArgGlyAsnIleValArgVal 286
QY 835 AGGGAGGCGCTCAGAAATCTTGAGCCAGATAGAAGAGACGTGCGCAGCAAGAGGAA 894
DB 287 LysGlyProPheGlnValValArgProLeuArgGln----- 299
QY 895 TACGATGAAGATGAATATACGATCAAGAGGATAGAAGCGTGGCGAGGGAAGCAGA 954
DB 300 -----ProTyrGluSerGluGlnTrpArgHisProArgGlyProProGlnSer 315
QY 955 GCGAGGGGAATGTGATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAGAACCAATTGCT 1014
DB 316 ProGlnAspAsnGlyLeuGluGluThrIleCysSerMetArgThrHisGluAsnIleAsp 335
QY 1015 AGAAACAGATCCCTGACATCTACACCCCTCAAGCTGCTTCACTCAAACTCCCAACGAT 1074
DB 336 AspProAlaArgAlaAspValTyrLysProAsnLeuGlyArgValThrSerValAsnSer 355
QY 1075 CTCACACTTCTAATACTAGTGGCTTGACCTAGTCTGTAATATGGAAATCTCTACAGG 1134
DB 356 TyrThrLeuProIleLeuGlnIleArgLeuSerAlaThrArgGlyIleLeuGlnGly 375
QY 1135 AATGATTGTTGTCGCTCACTACAAACACGACACAGCATCATATATCGATTGAGG 1194
DB 376 AsnAlaMetValLeuProLysTyrAsnMetAsnAlaAsnGluIleLeuTyrCysThrGln 395
QY 1195 GGACGGCTCAGTGCCTCAAGTCTGAGCAGCAGCAACGCGCAACAGAGTGTACGACGAGAGCT 1254
DB 396 GlyGlnAlaArgIleGlnValValAsnAspAsnGlyGlnAsnValLeuAspGlnGlnVal 415
QY 1255 CAAGAGGTGACGTGCTGTGTGTCACAGAACTTCGCGCTGCGTGGAAAGTCCAGAGC 1314
DB 416 GlnLysGlyGlnLeuValValIleProGlnGlyPheAlaTyrValValGlnSerHisGln 435
QY 1315 GAGAACTTCGAAATACGTGCGCATTCAGACAGACTCAAGGCCCGCAGCATAGCCAACTCGCC 1374
DB 436 AsnAsnPheGluTyrPheSerPheLysThrAsnAlaAsnAlaMetValSerThrLeuAla 455
QY 1375 GGTGAAAACCTCGCTCATAGATAAATCTGCGGAGGAGGTGTTGCAAAATTCATATGCGCTC 1434
DB 456 GlyArgThrSerAlaLeuArgAlaLeuProLeuGluValIleThrAsnAlaPheGlnIle 475
QY 1435 CAAGGGAGCAGGCAAGCGCAGCTTAAGAACAC 1467
DB 476 SerLeuGluGluAlaArgArgIleLysPheAsn 486
```

RESULT 12

```
US-10-508-263-66
; Sequence 66, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-508-263-66
```

Alignment Scores:
 Pred. No.: 1,34e-54 Length: 455
 Score: 734.50 Matches: 171
 Percent Similarity: 47.9% Conservative: 73
 Best Local Similarity: 33.6% Mismatches: 174
 Query Match: 26.4% Indels: 91
 DB: 9 Gaps: 8

US-10-728-323-3 (1-1524) x US-10-508-263-66 (1-455)

QY	19	AACGGCTCCAGATTCCAGCGCTCAATGCGCAGAGACCTTGACAATCGCATTCGAATTCGAATCAGAG	78
		:::	:::
DB	28	AsnGluCysGlnLeuAspGlnLeuAsnAlaLeuGluProSerGlnIleIleIysSerGlu	47
QY	79	GGCGGTTACATTGAGACTTGAACCCCAACCAACAGGAGTTCGAATGCGCGCGGTGCGC	138
DB	48	GlyGlyArgIleGluValTrpAspHisHisAlaProGlnLeuArgCysSerGlyPheAla	67
QY	139	CTCTCTGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCGCTTCTACTCCAAATGCTCC	198
DB	68	PheGluArgPheValIleGluProGlnGlyLeuPheLeuProThrPheLeuAsnAlaGly	87
QY	199	CAGGAGATCTTCATCCACCAAGGAGGGATACTTTGGGTGATATTCCTCGTTGTCTCT	258
DB	88	LysLeuThrPheValValHisGlyArgGlyLeuMetGlyArgValIleProGlyCysAla	107
QY	259	AGACACTATGAAGCGCTCACACACAAGTCTCGATCTCAGTCCCAAGACCAACAAGA	318
DB	108	GluThrPheMetGluSerProValPheGlyGluGly	119
QY	319	CGTCTCCAAGAGAAGACCAAGCCACAGCAAGATAGTCACCAAGGTGCACCGT	378
DB	120	-----GlnGlyGlnSerGlnGlyPheArgAspMetHisGlnIysValGluHis	137
QY	379	TTCGATGAGGTGATCTCATTCGAGTTCACCGGTGTGCTTTCTGGCTCTACAAACGAC	438
DB	138	LeuArgCysGlyAspThrIleAlaThrProSerGlyValIleGlnTrpPheTrpAsnAsn	157
QY	439	CACGACACTGATGTTGTTGCTGTTCTCTTACTGACACCAACAACACCAACAGCTT	498
DB	158	GlyAsnGluProLeuIleLeuValAlaAlaAspLeuAlaSerAsnGlnAsnGlnLeu	177
QY	499	GATCAGTTCCTCCAGGAGATTCAAATTTGGCTGGAAACACGGAG-----CAAGAGTTCTTTA	552
DB	178	AspArgAsnLeuArgProPheLeuIleAlaGlyAsnAsnProGlnGlyGlnGluTrpLeu	197
QY	553	AGGTACACAGCAACAAGCAGACAAGACGACGAAAGAGCTTACCATATAGCCCATACAGC	612
DB	198	GlnGlyArgIysGlnGlnIysGlnAsn-----	206
QY	613	CCGCAAGTTCAGCCTAGACAAGAAGCGTGAAATTTAGCCCTCGAGGACGACAGCCGC	672
DB	206	-----	206
QY	673	AGAGNACGAGCAGGACACAGAGAAGAAACGAGGTGGNAAACATCTTCAGCGCTTCAGC	732
DB	207	-----AsnIlePheAsnGlyPheAla	213
QY	733	CCGGAGTTCCTGGAAACAAGCCTTCCAG-----GTTGACGACAGACAGATAGTGCAAAAC	786
DB	214	ProGluIleLeuAlaGlnAlaPheIysIleAsnValGluThrAlaGlnGlnLeuGlnAsn	233
QY	787	CTAAGGCGGAGACCGAGAGTGAAGAAGGAGGCCATTGTGACAGTAGTAGGGGAGCCCTC	846
DB	234	-----GlnGlnAspAsnArgGlyAsnIleValIysValAsnGlyProPhe	248
QY	847	AGAACTCTTGAGCCACAGATAGAAGAAGACGTGCCGCAAGAAAGAAATACGATGAAGT	906
DB	249	GlyValIleArgPro-----	253
QY	907	GAATATGAATACGATGAAGAGGATAGAAGCGCTGCGAGGGAAGCAGAGGCAGG-----	960
DB	254	-----ProLeuArgArgGlyGluGlnGlnProHisGlu	266

Qy	79	GGCGGTTACATTGAGACTTGAACCCCAACACGAGGAGTTCGAATGCGCGCGGTGCC	138
Db	48	AlaGlyArgIleGluValTrpAspHisAlaProGlnLeuArgCysSerGlyValSer	67
Qy	139	CTCTCTCGCTTAGTCCCTCCGCGCAACGCGCTTCGTAGGCGCTTCTACTCTCAATGCTCCC	198
Db	68	PheValArgTyrIleIleGluSerIysGlyLeuTyrLeuProSerPhePheSerThrAla	87
Qy	199	CAGGAGATCTTCATCCAGCAAGGAGGATACTTTGGTTGCATATTC-----	246
Db	88	ArgLeuSerPheValAlaLysGlyGluGlyLeuMetGlyArgValValLeuCysAlaGlu	107
Qy	247	-----CCTGGTTGCTCAGACTATGAAGAGCCTCAC	279
Db	108	ThrPheGlnAspSerSerValPheGlnProSerGlyGlySerProPheGlyGluGlyGln	127
Qy	280	ACAAAGGTCTCGATCTCAGCTCCCAAGACCACCAAGACTCTCCCAAGGAA-----	333
Db	128	GlyGlnGlyGlnGlyGlnGlyGlnGlyHisGlnGlyGlnGlyGlnGlyGlnGly	147
Qy	334	-----GACCAAGCCCAACAGCAA-----CGAGATAGTCCACAGAAAG	369
Db	148	GlnGlnGlyGlnGlnGlnSerGlnGlyGlnGlyPheArgAspMetHisGlnLys	167
Qy	370	GTGCACCGTTTCGATGAGGGTGATCTCATTCAGGTCCACCGGTGTGCTTCTCGGCTC	429
Db	168	ValGluHisIleArgThrGlyAspThrIleAlaThrHisProGlyValAlaGlnTrpPhe	187
Qy	430	TACAACGACCAAGACACTGATGTTGTGCTGCTTTCTTCTTACTGACCAACCAACGAC	489
Db	188	TyrAsnAspGlyAsnGlnProLeuValIleValSerValLeuAspLeuAlaSerHisGln	207
Qy	490	AACCAGCTTGATCAGTTCCTCCAGGAGATTCAATTTGGCTGGGAACACGAGCAAGAGTTC	549
Db	208	AsnGlnLeuAspArgAsnProArgProPheTyrLeuAlaGlyAsn-----	222
Qy	550	TTAAGGTACCAGCAACAAGACAGCAACAAGCAGACGAAGAAGCTTACCATATAGCCCATAC	609
Db	222	-----	222
Qy	610	AGCCGGCAAGTCCAGCTAGACACAGAGAGCGTGAAATTAGCCCTCGAGGACAGACAGC	669
Db	223	AsnProGlnGlyGlnValTrpIleGluGlyArgGluGlnGlnProGln-----	238
Qy	670	CGCAGAAACGAGCAGGACAAAGAAGAAAAACGAAGGTGGAACATCTTCAGCGGCTTC	729
Db	239	-----LysAsnIleLeuAsnGlyPhe	245
Qy	730	ACGCCGAGTTCTCGGAACAAGCCCTCCAGTTGACGACAGA-----CAGATAGTGCAA	783
Db	246	ThrProGluValLeuAlaLysAlaPheLysIleAspValArgThrAlaGlnGlnLeuGln	265
Qy	784	AACCTAGAGCGGACACGAGAGTGAAGAGAGGAGCCATTGTGACAGTGAAGGGGAGGC	843
Db	266	Asn-----GlnGlnAspAsnArgGlyAsnIleIleArgValGlnGlyPro	280
Qy	844	CTCAGAATCTTGAGCCAGATAGAAAGAGACGTCCGACGAAAGAGGAATACGATCAA	903
Db	281	PheSerValIleArgProProLeuArgSerGlnArgProGlnGluGluVal-----	297
Qy	904	GATGAATATGAATACGATGAAGAGGATAGAAAGGCGTGGCAGGGGAAGCAGAGGAGGGGG	963
Db	297	-----	297
Qy	964	AATGGTATTGAAGACGATCTGCACCGCAAGTCTGCTAAAAGAACATTGGTAGAAACAGA	1023
Db	298	AsnGlyLeuGluGluThrIleCysSerAlaArgCysThrAspAsnLeuAspAspProSer	317
Qy	1024	TCCCTTCACATCTCAACCTTCAGCTGGTTCTCACTCAAACTGCCCAACGATCTCAACTT	1083
Db	318	AsnAlaAspValTyrLysProGlnLeuGlyTyrIleSerThrLeuAsnSerTyrAspLeu	337
Qy	1084	CTAATACTTAGGTGGCTTGGACCTTAGTGTCTGAATATGAAATCTCTACAGGAATGCATTG	1143

[illegible]

```
Db      108  GluThrPheGlnAspSerValPheGlnProGlySerGlySerProPheGlyGluGly 127
QY      277  CACACACAAGGTCGTGATCTCAG-----TCC 303
Db      128  GlnGlyGlnGlnGlyGlnGlyGlnGlyGlnGlyGlnGlyGlnGlyGlnGly 147
QY      304  CAAAGACCACCAAGAGCGTCTCAAGAGGAGAAGACCAAGACCAACAGCAA-----351
Db      148  GlnGlyGlnGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGly 167
QY      352  ---CGAGATAGTCACACAGAGGTGCACCCGTTTCGATGAGGGGTGATCTCAATTGCAGTCCC 408
Db      168  PheArgAspMetHisGlnLysValGluHisIleArgSerGlyAspThrIleAlaThrHis 187
QY      409  ACCGGTGTTCCTTTCGGCTGTACAAAGACACGACACATGATGTGTGCTGTTCTCTTT 468
Db      188  ProGlyValAlaGlnTrpPheTyrAsnAsnGlyAsnGlnProLeuValIleValAlaVal 207
QY      469  ACTGACACCAACAACAACACACAGCTTGATCAGTTCGCCAGGAGATTCAATTTGGCT 528
Db      208  MetAspLeuAlaSerHisGlnAsnGlnLeuAspArgAsnProSerGlyPheTyrLeuAla 227
QY      529  GGGAAACACGGAGCAAGATTCTTAAGGTACCAGCAACAAGACAGACAAAGACAGCAAGA 588
Db      228  GlyLys-----229
QY      589  AGCTTACCATATAGCCCATACAGCCCGCAAGTCAGCCTAGACAAAGAGCGTGATTT 648
Db      229  -----229
QY      649  AGCCCTCGAGGACAG-----CACAGCGCAGAGACGAGCAGCAAGAAGAAGAA 699
Db      230  AsnProGlnGlyGlnSerTrpLeuHisGlyArg-----GlyGlnGlnProGln 245
QY      700  AACGAAGGTGGAACAATCTTCAGCGGCTTCACGCCGAGTTCCTGGAACAAGCCTTCCAG 759
Db      246  Asn-----AsnIleLeuAsnGlyPheSerProGluValLeuAlaGlnAlaPheLys 262
QY      760  GTTGACGACAGA-----CAGATAGTGCAAAACCTAAGAGGCGAGACCGAGAGTGAAGAA 813
Db      263  IleAspValArgThrAlaGlnGlnLeuGlnAsn-----GlnGlnAspAsn 277
QY      814  GAGGAGGACCATCTGACAGTGGAGGAGGCGCTCAGAAATCTTGAGCCCGACATAGAAAGAGA 873
Db      278  ArgGlyAsnIleValArgValGlnGlyProPheGlyValIleArgProLeuLysSer 297
QY      874  CGTCCGACGAGAAGAAGAAATACGATGAAGATGAATATCAATACGATGAAGAGGATAGA 933
Db      298  GlnArgProGlnGluThrGlu-----304
QY      934  AGCGGTGGCAGGGAAGCAGAGCAGGGGGAATGGTATTGAAAGACGATCTGCACCGCA 993
Db      305  -----AlaAsnGlyLeuGluGluThrIleCysSerAla 315
QY      994  AGTGCTAAAAAGAACATTGTGTAGAAACAGATCCCTTGACATCTACAAACCTCAAGCTGT 1053
Db      316  ArgCysThrAspAsnLeuAspAspProSerAsnAlaAspValTyrLysProGlnLeuGly 335
QY      1054  TCACCTCAAACTGCCAACGATCTCAACCTTCTAATATCTAGTGGCTGGACCTAGTGCT 1113
Db      336  TyrIleSerIleLeuAsnSerTyrAspLeuProIleLeuArgValLeuArgLeuSerAla 355
QY      1114  GAATATGGAAATCTTACAGGAATGATGTTTGTGCTCCTCACTACAAACCAACGACACAC 1173
Db      356  LeuArgGlySerIleArgGlnAsnAlaMetValLeuProGlnTrpLysSerTyrSerAsn 375
QY      1174  AGCATCATATATCGATTGAGGGACGGGCTCAGTGCAGTCTGTGACGACAGCAAC 1233
Db      376  AlaValIleTyrValThrAspGlyGluAlaGlnIleGlnValValAsnAspAsnGlyAsp 395
QY      1234  AGAGTGTACGACGAGAGCTTCAAGAGGGTTCAGTGTGTGTGTCGCCACAGAACTTCGCC 1293
Db      396  ArgValPheAspGlyGlnValSerGlnGlnLeuLeuSerIleProGlnGlyPheSer 415
```

```
QY      1294  GTCGCTGGAAGTCCCGAGCGAGAACTTCGATACGTGGCATTCAAGACAGACTCAAGG 1353
Db      416  ValValLysArgAlaThrSerAspGlnPheArgTrpIleGluPheLysThrAsnAlaAsn 435
QY      1354  CCCAGCATAGCCCAACTCGCCGTTGAAAACTCCGTCATAGATAAACCCTGCCGAGGAGGTG 1413
Db      436  AlaGlnIleAsnThrLeuAlaGlyArgThrSerValMetArgGlyLeuProLeuGluVal 455
QY      1414  GTTGCAAAATTCATATGGCTCCAAAGGAGGAGCGAGGCAAGCAGCTTAAGAACAAAC 1467
Db      456  IleAlaAsnGlyTyrGlnIleSerLeuGluGluAlaArgArgValLysPheAsn 473
```

RESULT 15

```
US-10-508-263-68
; Sequence 68, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 68
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-508-263-68
```

```
Alignment Scores:
Pred. No.:          9,4e-51      Length:      451
Score:              689.50      Matches:    157
Percent Similarity: 48.0%      Conservative: 85
Best Local Similarity: 31.2%   Mismatches: 174
Query Match:        24.8%      Indels:     88
DB:                  9          Gaps:         8
```

US-10-728-323-3 (1-1524) x US-10-508-263-68 (1-451)

```
QY      1  CGCGCAGCAACCGGAGGAGAACCGTCGCCAGTTCCAGCGCTCAATCGCGAGACCTGAC 60
Db      28  ArgGluAlaProPheProAsnAlaCysHisPheSerGlnIleAsnSerLeuAlaProAla 47
QY      61  AATCGCATTAATCATGAGCGCGTTACATTGAGACTTGGAACCCCAACACGAGGAGTTTC 120
Db      48  GlnAlaThrLysPheGluAlaGlyGlnMetGluValTrpAspHisMetSerProGluLeu 67
QY      121  GAATGCGCGCGTCGCCCTCTCTCGTTAGTCTCCGCGCAACCGCCTTCGTAGGCCT 180
Db      68  ArgCysAlaGlyValThrValAlaArgIleThrLeuGlnProAsnSerIlePheLeuPro 87
QY      181  TTCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGGGATCTTTGGGTG 240
Db      88  AlaPhePheSerProProAlaLeuAlaTyrValValGlnGlyGluGlyValMetGlyThr 107
QY      241  ATATTCCCTGGTTGCTCTAGACACTATGAAAGACCTCACACACAGGTCGTGCATCTCAG 300
Db      108  IleAlaSerGlyCysProGluThrPheAlaGluValGluGlySerSerGlyArgGlyGly 127
QY      301  TCCCAAGACCAACCAAGACGCTCTCCAAGGAGAGAAGCAAAAGCCACAGCAACGAGTAGT 360
Db      128  GlyGlyAspProGlyArgArgPheGlu-----AspMet 138
QY      361  CACCAAGGTGCACCGTTTCGATGAGGGTGATCTCATTCGATTCCTCCACCGGTGTGCT 420
Db      139  HisGlnLysLeuGluAsnPheArgArgGlyAspValPheAlaSerLeuAlaGlyValSer 158
QY      421  TTCTGGCTCTACAAACCAACCAACGACTGATGTGTGTGTTCTCTTCTTACTGACCAAC 480
Db      159  GlnTrpTrpTyrAsnArgGlyAspSerAspAlaValIleValLeuAspValThr 178
```


THIS PAGE LEFT BLANK

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame plus n2p model

Run on: May 15, 2006, 21:46:12 ; Search time 10.2004 Seconds
(without alignments)
4312.597 Million cell upd

Title: US-10-728-323-3

Perfect score:

Sequence: 1 cggcagcaaccggaggagaa.....ctccgagggctgtggcttaa 1524

Scoring table:

Xgapop	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Fgapop	6.0	Fgapext	7.0
Delop	6.0	Delext	7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Minimum DB seq length:	0
Maximum DB seq length:	2000000000

Post-processing: Minimum

Post-processing: Minimum Match 0%
Maximum Match 100%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ n2p.model -DEV=xlp
```

-Q=/abgs/ABSSWEB spool/US10728323/runat 15052006 172134 22413/app query.fasta 1

```
-DB=PIR -QFMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEX
```

```
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
```

```
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
```

```
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss02p
```

```
-USER=US10728323 @CGN_1_1_92 @runat_15052006_172134_22413 -NCPU=6 -ICPU=3
```

```
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
```

```
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
```

```
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database : PIR 80:*

```
1: _pir1:*
```

```
2: pir2:*
```

```
3: pir3:*
```

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	1502	54.0	481	2	S04605	glycinin G3 - soyb
2	1499.5	54.0	485	2	S11002	glycinin G2 precu
3	1496	53.8	495	2	S10851	glycinin G1 precu
4	1492.5	53.7	485	1	FWSVG1	glycinin chain A2B
5	1466	52.8	495	1	FWSVG2	glycinin chain A1a
6	1457.5	52.4	498	2	S04394	legumin A2 precursor
7	1455	52.4	520	2	S08237	legumin A2 precursor
8	1440	51.8	517	1	FWPMLA	legumin A precursor
9	1431.5	51.5	482	2	S49877	legumin A precursor
10	1428	51.4	484	2	S11003	glycinin G3 precu
11	1414.5	50.9	500	2	S14393	legumin A2 precursor
12	1384	49.8	507	2	T06452	probable legumin A
13	1375	49.5	497	2	S14392	legumin A1 precursor
14	1073.5	38.6	551	2	S51941	prunin 1 precursor

Alignment Scores:

Assignment Score:	1.34e-103	Length:	481
Pred. No.:	1502.00	Matches:	237
Score:	71.6%	Conservative:	68
Percent Similarity:	58.2%	Mismatches:	91
Best Local Similarity:	58.2%	Indels:	54
Query Match:	54.0%	Gaps:	7
DR:	2		

US-10-728-323-3 (1-1524) x 504605 (1-481)

Qy	1	CGG	CAG	CAC	CGG	AGG	AG	A	CG	CGT	GCT	CAG	CGC	CTC	AAT	GCG	CAG	AGC	CTG	AC	60	
Db	23	Arg	Glu	Gln	Pro	Gln	Asn	Glu	Cys	Gln	Leu	Gln	Arg	Leu	Asn	Ala	Leu	Val	Pro	Asp	42	
Qy	61	AAT	CGC	ATT	GAA	T	CAG	AGG	CGG	TTC	A	ATT	GAG	ACT	TGG	AAC	CCC	CAA	CAC	CAG	AGTTC	120
Db	43	Asn	Arg	Ile	Glu	Ser	Glu	Gly	Gly	Phe	Ile	Glu	Thr	Trp	Asn	Pro	Asn	Asn	Val	Pro	Phe	62
Qy	121	GAAT	GCG	CGG	CGT	CGC	CTC	TCT	GCT	TAG	TCT	CTC	CGC	CGA	AGC	CC	CTT	CGT	AGC	CT	180	

Db 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 82
:::|||||
QY 181 TTCTACTCCAAATCCTCCAGAGATCTTCATCCAGCAGGAGGAGTACTTTGGTTG 240
|||||
Db 83 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 102
:::|||||
QY 241 ATATTCCTGTTGTCCTAGACACTATGAAGACCTTCACACACAAGGTCGTGCATCTCAG 300
|||||
Db 103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnGly 119
:::|||||
QY 301 TCCCAAGACCAACCAAGCGTCTCCAGGAGAGACCAAGCCACAGCAAGAGATAGT 360
|||
Db 120 SerSerArgPro 126
:::|||||
QY 361 CACCAAGGTGCACCGTTTCGATGAGGTTGATCTCATTCGAGTTCCTCCACCGTGTGCT 420
|||||
Db 127 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146
:::|||||
QY 421 TTCTGCTCTAACACGACGACACTGATGTTGTTGCTGTTCTTCTTACTGACCAAC 480
|||||
Db 147 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 166
:::|||||
QY 481 AACACGACCAACAGCTTGATCAGTTCCTCCAGGAGATTCAATTTGGCTGGACACGGAG 540
|||||
Db 167 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
:::|||||
QY 541 CAAGAGTTCTTAAAGTACCAACCAAGCAGACAGCAAGAGAGAGAGAGAGAGAGAGAG 600
|||||
Db 187 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly 199
:::|||||
QY 601 AGCCCATACAGCCCGCAAGTCAAGCTAGACAAAGAGAGCGGTGAATTTAGCCCTCAGGA 660
|||
Db 200 -----Gly 200
:::|||||
QY 661 CACACAGCGCAGAGACGAGCAGGACGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
|||||
Db 201 ThrGlnSerGlnLysGlyLysArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
:::|||||
QY 721 AGCGGCTTCACGCGCGGAGTTCTTGGAAACAAAGCCTTCAGGTTGACGACAGACAGATAGT 780
|||||
Db 221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 239
:::|||||
QY 781 CAAACCTAAGAGCGGACGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
:::|||||
Db 240 ArgLysLeuGlnGlyGluAsnGluGluGluGluGlyGlyAlaIleValThrValLysGly 259
|||||
QY 841 GGCCTCAGATCTTGAGCCCA-----GATAGAAAGAGAGCTGCCGACGAGAGAGAG 891
|||||
Db 260 GlyLeuSerValIleSerProProThrGluGluGlnGlnGlnArgProGluGluGluGlu 279
:::|||||
QY 892 GAATACGATGAAGATGAATATGAATACGATGAAGAGATAGAAAGCGCTGCCGAGGAGC 951
:::|||||
Db 280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294
|||
QY 952 AGAGGAGGCGGGAATGGTATTGAAGAGAGATCTGCACCGCAAGTCTAAAGAGACATT 1011
|||||
Db 295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311
:::|||||
QY 1012 GGTAGAAACAGATCCCTGCATCTCAACCCCTCAAGCTGGTTCACCTCAAAACCTGCCAAC 1071
|||||
Db 312 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrAlaThr 331
:::|||||
QY 1072 GATCTCAACCTCTTAATCTAGGTGGCTGGACCTAGTGTGAATATGAAATCTCTAC 1131
|||||
Db 332 SerLeuAspPheProAlaLeuSerTyrLeuLysLeuSerAlaGlnPheGlySerLeuArg 351
:::|||||
QY 1132 AGAATGCACTGTTTGTGCTCTACTCAACACCAACGACGACATCATATATCATGTC 1191
:::|||||
Db 352 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 371
:::|||||
QY 1192 AGGGGAGCGGCTCACGTGCAAGTCGTGGACAGCAACGGCAACAGAGTGTACGACGAGGAG 1251
|||||

Db 372 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 391
QY 1252 CTTCAAGAGGGTCACGTGCTTGTGTGTCACAGAACTTCGCGCTGCGTGAAGATCCCGAG 1311
|||||
Db 392 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 411
:::|||||
QY 1312 AGCGAGAACTTCGATAGCTGATTCAGACAGACTCAAGGCCCGACATAGCCACCTC 1371
|||||
Db 412 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431
:::|||||
QY 1372 GCGCGTGAACCTCCGTCATAGATAACCTGCCGAGAGAGGTGGTTCCAAATTTCATATGCG 1431
|||||
Db 432 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnThrPheAsn 451
:::|||||
QY 1432 CTTCAAGAGGACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1491
|||||
Db 452 LeuArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 471
:::|||||
QY 1492 CCGTCTCAGCAGCTCTCCGAGGCGCTGCGCT 1521
|||
Db 472 ProLysGluSerGlnArgValValAla 481
|||
RESULT 2
S11002
glycinin G2 precursor - soybean
N:Alternate names: glycinin A2B1a
C:Species: Glycine max (soybean)
C:Date: 21-Nov-1993 #sequence revision 19-Jan-1996 #text_change 05-Oct-2004
C:Accession: S11002; S04604; A26990
R:Nielsen, N.C.; Dickinson, C.D.; Cho, T.J.; Thanh, V.H.; Scallion, B.J.; Fischer, R.L.;
Plant Cell 1, 313-328, 1989
A:Title: Characterization of the glycinin gene family in soybean.
A:Reference number: S10851; MUID:92393391; PMID:2485233
A:Accession: S11002
A:Molecule type: DNA
A:Residues: 1-485 <NIE>
A:Cross-references: UNIPROT:P04405; UNIPARC:UPI000012B557
A:Experimental source: variety bare
R:Thanh, V.H.; Tumer, N.E.; Nielsen, N.C.
Nucleic Acids Res. 17, 4387, 1989
A:Title: The glycinin Gy(2) gene from soybean.
A:Reference number: S04604; MUID:89296499; PMID:2740230
A:Accession: S04604
A:Molecule type: DNA
A:Residues: 1-485 <THA>
A:Cross-references: UNIPARC:UPI000012B557; EMBL:X15122; NID:g18636; PIDN:CAA33216.1; PI:
A:Experimental source: variety bare
R:Fukazawa, C.; Momma, T.; Higuchi, W.; Udaka, K.
Nucleic Acids Res. 15, 8117, 1987
A:Title: Complete nucleotide sequence of the gene encoding a glycinin A2B1a subunit pre
A:Reference number: A26990; MUID:88040439; PMID:3671077
A:Accession: A26990
A:Molecule type: DNA
A:Residues: 1-485 <PUK>
A:Cross-references: UNIPARC:UPI000012B557; GB:X02806
C:Genetics:
A:Gene: GY2
A:Introns: 93/1; 177/3; 356/3
C:Keywords: storage protein
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-485/Product: glycinin G2 #status predicted <MAT>
Alignment Scores:
Pred. No.: 2.05e-103 Length: 485
Score: 1499.50 Matches: 297
Percent Similarity: 72.0% Conservative: 71
Best Local Similarity: 58.1% Mismatches: 94
Query Match: 54.0% Indels: 49
DB: 2 Gaps: 6
US-10-728-323-3 (1-1524) x S11002 (1-485)
QY 1 CGGAGCAACCGAGGAGAGAACGCGTGCCAGGTTCCAGCGCTCAATGCGCAGACCTGAC 60

Alignment Scores:				
Pred. No.:	3,74e-103	Length:	495	
Score:	1496.00	Matches:	292	
Percent Similarity:	70.9%	Conservative:	73	
Best Local Similarity:	56.7%	Mismatches:	100	
Query Match:	53.8%	Indels:	50	
DB:	2	Gaps:	6	
US-10-728-323-3 (1-1524) x S10851 (1-495)				
QY	1	CGGCAGCAACCGGAGGAGAACGGCTGCCAGTTCACAGCGCTCAATGCGCAGAGACCTGAC	60	
DB	23	ArgGluGluProGlnGlnAnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProAsp	42	
QY	61	AATCGCATTCGAATCAGAGCGCGTTTACATTGAGACTTGGAACCCCAACACAGGAGTTTC	120	
DB	43	AsnArgIleGluSerGluGlyGlyLeuIleGluThrTrpAsnProAsnAlaLeuArgProPhe	62	
QY	121	GAATGCGCGGCGTCCGCTCTCTCGCTTAGTCTTCGCGCGCAACGCCCTTCGTAGGCGCT	180	
DB	63	GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro	82	
QY	181	TTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGGGATACTTTGGGTTG	240	
DB	83	SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyLysIlePheGlyMet	102	
QY	241	ATATTCCTCGTGTCTCTAGACACTATGAAGAGCGCTCACACAAAGGTCGTGATCTCAG	300	
DB	103	IleTyrProGlyCysProSerThrPheGluGluProGlnGlnProGlnArgGlyGln	122	
QY	301	TCCCAAGACCCACCAAGAGCTCTCCAGGAGAGACCAAGCCACAGCAACAGATAGT	360	
DB	123	SerSerArgPro-----GlnAspArg	129	
QY	361	CACCAGAGTGCACCGTTTCGATGAGGGTGATCTCATTTGCAGTTCACCACCGGTGTGCT	420	
DB	130	HisGlnLysIleTyrAsnPheArgGluGlyAspLeuIleAlaValProThrGlyValAla	149	
QY	421	TTCTGCTCTTACACGACACCGACACTGATGTTGTTGCTGTCTTCTTACTGACACCAAC	480	
DB	150	TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerIleIleAspThrAsn	169	
QY	481	AACAACGACACACAGCTTGATCAGTTCCTCCCGAGAGATTCAATTTGGCTGGGACACGGAG	540	
DB	170	SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu	189	
QY	541	CAAGAGTTCTTAAGGTACAGCAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	600	
DB	190	GlnGluPheLeuLysTyrGlnGlnGluGln-----	199	
QY	601	AGCCCATACAGCCCGCAAAAGTTCAGCCTAGACAAAGAGAGCGTGAAATTTAGCCCTCGAGGA	660	
DB	200	-----GlyGly	201	
QY	661	CACACAGCGCGCAGAGAACGACGAGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG	720	
DB	202	HisGlnSerGlnLysGlyLysHisGlnGlnGluGluAsnGluGlyGlySerIleLeu	221	
QY	721	AGCGGCTTACGCGCGGAGTTCCTGGAAACAGCGCTTCAGGTTTCAGCAGACAGATAGTG	780	
DB	222	SerGlyPheThrLeuGluPheLeuGluHisAlaPheSerVal---AspLysGlnIleAla	240	
QY	781	CAAAACCTAAGAGCGGACCGGAGAGTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	840	
DB	241	LysAsnLeuGlnGlyGluAsnGluGlyGluAspLysGlyAlaIleValThrValLysGly	260	
QY	841	GGCTCAGATCTTGAGCCCA-----GATAGAAGAGAGCTCCGACGAGAGAGAG	891	
DB	261	GlyLeuSerValIleLysProProThrAspGluGlnGlnArgProGlnGluGluGlu	280	
QY	892	GAATACGATCAAGATGAATGAATAC-----GATGAAGAG-----GATGAAGG	936	

DB	281	GluGluGluGluAspGluLysProGlnCysLysGlyLysAspLysHisCysGlnArgPro	300	
QY	937	CGTGCAGGGGAAGCAGAGCGAGGGGAATGTATTGAACAGACGATCTGCCACCGCAAGT	996	
DB	301	ArgGlySerGlnSerLysSerArgAsnGlyIleAspGluThrIleCysThrMetArg	320	
QY	997	GCTAAAAAGAACATTGGTAGAAAACAGATCCCTCGATCATCTACAACCTCAAGCTGGTTCA	1056	
DB	321	LeuArgHisAsnIleGlyGlnThrSerSerProAspIleTyrAsnProGlnAlaGlySer	340	
QY	1057	CTCAAAACTGCCAACGATCTCAACCTTCTTAATCTAGTGGCTGGACCTAGTGTGAA	1116	
DB	341	ValThrThrAlaThrSerLeuAspPheProAlaLeuSerTrpLeuArgLeuSerAlaGlu	360	
QY	1117	TATGMAATCTCTACAGGAATGCATTGTTGTCGCTCACTACAACCAACGACACACAGC	1176	
DB	361	PheGlySerLeuArgLysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSer	380	
QY	1177	ATCATATATTCGATTGAGGGGACGGCTCACGTGCAAGTCTGTGACAGCAACGGCAACAGA	1236	
DB	381	IleIleTyrAlaLeuAsnGlyArgAlaLeuIleGlnValValAsnCysAsnGlyGluArg	400	
QY	1237	GTGTACGAGAGAGCTTCAAGAGGGTTCACGTGCTTGTGTCGCCACAGAACTTCGCCGT	1296	
DB	401	ValPheAspGlyLeuLeuGlnGluGlyArgValLeuIleValProGlnAsnPheValVal	420	
QY	1297	GCTGMAATCCCGACGAGCGAGAACTTCGAATACGTGGCATCTCAAGACACTCAAGGCC	1356	
DB	421	AlaAlaArgSerGlnSerAspAsnPheGluTyrValSerPheLysThrAsnAspThrPro	440	
QY	1357	AGCATAGCCAACTCGCGGTGAAACTCCGTCTATAGATAAACCCTGCGGAGGAGTGGTT	1416	
DB	441	MetIleGlyThrLeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIle	460	
QY	1417	GCAAAATTCATATGCTCCAAAGGAGGAGCGAAGCAAGCACTTAAGAACAAACCCCTTC	1476	
DB	461	GlnHisThrPheAsnLeuLysSerGlnGlnAlaArgGlnIleLysAsnAsnProPhe	480	
QY	1477	AGTTCCTTCGTTCCACCGCTCTCAGCAGTCTCCGAGGGCTGTGGCT	1521	
DB	481	LysPheLeuValProProGlnGluSerGlnLysArgAlaValAla	495	
RESULT 4				
FWSYGI				
glycinin chain A2B1a precursor - soybean				
N;Alternate names: 11S globulin				
N;Contains: Glycinin chain A2; glycinin chain B1a				
C;Species: Glycine max (soybean)				
C;Date: 31-Mar-1988 #sequence revision 31-Mar-1988 #text change 05-Oct-2004				
C;Accession: A91341; A92454; A90024; A90024; A92452; S10503; S74123; A05082; A05164; A0				
R;Mamma, T.; Negro, T.; Uda, K.; Fukazawa, C.				
FBS Lett. 188, 117-122, 1985				
A;Title: A complete cDNA coding for the sequence of glycinin A2B1a subunit precursor.				
A;Reference number: A91341				
A;Accession: A91341				
A;Molecule type: mRNA				
A;Residues: 1-485 <MOM>				
A;Cross-references: UNIPROT:P04405; UNIPARC:UPI000016DD98				
A;Experimental source: strain Bonimori				
A;Note: the source of this mRNA was cotyledon tissue taken from seeds at the middle sta				
R;Marco, Y.A.; Thanh, V.H.; Tumer, N.E.; Scallan, B.J.; Nielsen, N.C.				
J. Biol. Chem. 259, 13436-13441, 1984				
A;Title: Cloning and structural analysis of DNA encoding an A2B1a subunit of glycinin.				
A;Reference number: A92454; MUID:85030472; PMID:6092376				
A;Molecule type: mRNA				
A;Residues: 262-446 <MA1>				
A;Cross-references: UNIPARC:UPI00001745FF				
A;Accession: B92454				
A;Molecule type: DNA				
A;Residues: 318-485 <MA2>				
A;Cross-references: UNIPARC:UPI00001745FF				
A;Experimental source: strain CX635-1-1-1				

R:Utsumi, S.; Kim, C.S.; Kohno, M.; Kito, M.
 Agric. Biol. Chem. 51, 3267-3273, 1987
 A:Title: Polymorphism and expression of cDNAs encoding glycine subunits.
 A:Reference number: A90024
 A:Accession: A90024
 A:Molecule type: mRNA
 A:Residues: 1-38, 'D', '40-485 <UTS>
 A:Cross-references: UNIPARC:UPI000012B557
 A:Experimental source: strain Shirotsubunoko
 R:Staswick, P.E.; Hermodson, M.A.; Nielsen, N.C.
 J. Biol. Chem. 259, 13424-13430, 1984
 A:Title: The amino acid sequence of the A2B1a subunit of glycine.
 A:Reference number: A92452; MUID:85030470; PMID:6541652
 A:Accession: A92452
 A:Molecule type: protein
 A:Residues: 19-38, 'D', '40-60, 'S', '62-116, 'C', '118-192, 'E', '194-296, '301-342, 'S', '344-463, 'I', '465-485, 'D', '487-488, 'S', '490-491, 'D', '493-494, 'S', '496-497, 'D', '499-500, 'S', '502-503, 'D', '505-506, 'S', '508-509, 'D', '511-512, 'S', '514-515, 'D', '517-518, 'S', '520-521, 'D', '523-524, 'S', '526-527, 'D', '529-530, 'S', '532-533, 'D', '535-536, 'S', '538-539, 'D', '541-542, 'S', '544-545, 'D', '547-548, 'S', '550-551, 'D', '553-554, 'S', '556-557, 'D', '559-560, 'S', '562-563, 'D', '565-566, 'S', '568-569, 'D', '571-572, 'S', '574-575, 'D', '577-578, 'S', '580-581, 'D', '583-584, 'S', '586-587, 'D', '589-590, 'S', '592-593, 'D', '595-596, 'S', '598-599, 'D', '601-602, 'S', '604-605, 'D', '607-608, 'S', '610-611, 'D', '613-614, 'S', '616-617, 'D', '619-620, 'S', '622-623, 'D', '625-626, 'S', '628-629, 'D', '631-632, 'S', '634-635, 'D', '637-638, 'S', '640-641, 'D', '643-644, 'S', '646-647, 'D', '649-650, 'S', '652-653, 'D', '655-656, 'S', '658-659, 'D', '661-662, 'S', '664-665, 'D', '667-668, 'S', '670-671, 'D', '673-674, 'S', '676-677, 'D', '679-680, 'S', '682-683, 'D', '685-686, 'S', '688-689, 'D', '691-692, 'S', '694-695, 'D', '697-698, 'S', '700-701, 'D', '703-704, 'S', '706-707, 'D', '709-710, 'S', '712-713, 'D', '715-716, 'S', '718-719, 'D', '721-722, 'S', '724-725, 'D', '727-728, 'S', '730-731, 'D', '733-734, 'S', '736-737, 'D', '739-740, 'S', '742-743, 'D', '745-746, 'S', '748-749, 'D', '751-752, 'S', '754-755, 'D', '757-758, 'S', '760-761, 'D', '763-764, 'S', '766-767, 'D', '769-770, 'S', '772-773, 'D', '775-776, 'S', '778-779, 'D', '781-782, 'S', '784-785, 'D', '787-788, 'S', '790-791, 'D', '793-794, 'S', '796-797, 'D', '799-800, 'S', '802-803, 'D', '805-806, 'S', '808-809, 'D', '811-812, 'S', '814-815, 'D', '817-818, 'S', '820-821, 'D', '823-824, 'S', '826-827, 'D', '829-830, 'S', '832-833, 'D', '835-836, 'S', '838-839, 'D', '841-842, 'S', '844-845, 'D', '847-848, 'S', '850-851, 'D', '853-854, 'S', '856-857, 'D', '859-860, 'S', '862-863, 'D', '865-866, 'S', '868-869, 'D', '871-872, 'S', '874-875, 'D', '877-878, 'S', '880-881, 'D', '883-884, 'S', '886-887, 'D', '889-890, 'S', '892-893, 'D', '895-896, 'S', '898-899, 'D', '901-902, 'S', '904-905, 'D', '907-908, 'S', '910-911, 'D', '913-914, 'S', '916-917, 'D', '919-920, 'S', '922-923, 'D', '925-926, 'S', '928-929, 'D', '931-932, 'S', '934-935, 'D', '937-938, 'S', '940-941, 'D', '943-944, 'S', '946-947, 'D', '949-950, 'S', '952-953, 'D', '955-956, 'S', '958-959, 'D', '961-962, 'S', '964-965, 'D', '967-968, 'S', '970-971, 'D', '973-974, 'S', '976-977, 'D', '979-980, 'S', '982-983, 'D', '985-986, 'S', '988-989, 'D', '991-992, 'S', '994-995, 'D', '997-998, 'S', '999-1000, 'S', '1001-1002, 'S', '1003-1004, 'S', '1005-1006, 'S', '1007-1008, 'S', '1009-1010, 'S', '1011-1012, 'S', '1013-1014, 'S', '1015-1016, 'S', '1017-1018, 'S', '1019-1020, 'S', '1021-1022, 'S', '1023-1024, 'S', '1025-1026, 'S', '1027-1028, 'S', '1029-1030, 'S', '1031-1032, 'S', '1033-1034, 'S', '1035-1036, 'S', '1037-1038, 'S', '1039-1040, 'S', '1041-1042, 'S', '1043-1044, 'S', '1045-1046, 'S', '1047-1048, 'S', '1049-1050, 'S', '1051-1052, 'S', '1053-1054, 'S', '1055-1056, 'S', '1057-1058, 'S', '1059-1060, 'S', '1061-1062, 'S', '1063-1064, 'S', '1065-1066, 'S', '1067-1068, 'S', '1069-1070, 'S', '1071-1072, 'S', '1073-1074, 'S', '1075-1076, 'S', '1077-1078, 'S', '1079-1080, 'S', '1081-1082, 'S', '1083-1084, 'S', '1085-1086, 'S', '1087-1088, 'S', '1089-1090, 'S', '1091-1092, 'S', '1093-1094, 'S', '1095-1096, 'S', '1097-1098, 'S', '1099-1100, 'S', '1101-1102, 'S', '1103-1104, 'S', '1105-1106, 'S', '1107-1108, 'S', '1109-1110, 'S', '1111-1112, 'S', '1113-1114, 'S', '1115-1116, 'S', '1117-1118, 'S', '1119-1120, 'S', '1121-1122, 'S', '1123-1124, 'S', '1125-1126, 'S', '1127-1128, 'S', '1129-1130, 'S', '1131-1132, 'S', '1133-1134, 'S', '1135-1136, 'S', '1137-1138, 'S', '1139-1140, 'S', '1141-1142, 'S', '1143-1144, 'S', '1145-1146, 'S', '1147-1148, 'S', '1149-1150, 'S', '1151-1152, 'S', '1153-1154, 'S', '1155-1156, 'S', '1157-1158, 'S', '1159-1160, 'S', '1161-1162, 'S', '1163-1164, 'S', '1165-1166, 'S', '1167-1168, 'S', '1169-1170, 'S', '1171-1172, 'S', '1173-1174, 'S', '1175-1176, 'S', '1177-1178, 'S', '1179-1180, 'S', '1181-1182, 'S', '1183-1184, 'S', '1185-1186, 'S', '1187-1188, 'S', '1189-1190, 'S', '1191-1192, 'S', '1193-1194, 'S', '1195-1196, 'S', '1197-1198, 'S', '1199-1200, 'S', '1201-1202, 'S', '1203-1204, 'S', '1205-1206, 'S', '1207-1208, 'S', '1209-1210, 'S', '1211-1212, 'S', '1213-1214, 'S', '1215-1216, 'S', '1217-1218, 'S', '1219-1220, 'S', '1221-1222, 'S', '1223-1224, 'S', '1225-1226, 'S', '1227-1228, 'S', '1229-1230, 'S', '1231-1232, 'S', '1233-1234, 'S', '1235-1236, 'S', '1237-1238, 'S', '1239-1240, 'S', '1241-1242, 'S', '1243-1244, 'S', '1245-1246, 'S', '1247-1248, 'S', '1249-1250, 'S', '1251-1252, 'S', '1253-1254, 'S', '1255-1256, 'S', '1257-1258, 'S', '1259-1260, 'S', '1261-1262, 'S', '1263-1264, 'S', '1265-1266, 'S', '1267-1268, 'S', '1269-1270, 'S', '1271-1272, 'S', '1273-1274, 'S', '1275-1276, 'S', '1277-1278, 'S', '1279-1280, 'S', '1281-1282, 'S', '1283-1284, 'S', '1285-1286, 'S', '1287-1288, 'S', '1289-1290, 'S', '1291-1292, 'S', '1293-1294, 'S', '1295-1296, 'S', '1297-1298, 'S', '1299-1300, 'S', '1301-1302, 'S', '1303-1304, 'S', '1305-1306, 'S', '1307-1308, 'S', '1309-1310, 'S', '1311-1312, 'S', '1313-1314, 'S', '1315-1316, 'S', '1317-1318, 'S', '1319-1320, 'S', '1321-1322, 'S', '1323-1324, 'S', '1325-1326, 'S', '1327-1328, 'S', '1329-1330, 'S', '1331-1332, 'S', '1333-1334, 'S', '1335-1336, 'S', '1337-1338, 'S', '1339-1340, 'S', '1341-1342, 'S', '1343-1344, 'S', '1345-1346, 'S', '1347-1348, 'S', '1349-1350, 'S', '1351-1352, 'S', '1353-1354, 'S', '1355-1356, 'S', '1357-1358, 'S', '1359-1360, 'S', '1361-1362, 'S', '1363-1364, 'S', '1365-1366, 'S', '1367-1368, 'S', '1369-1370, 'S', '1371-1372, 'S', '1373-1374, 'S', '1375-1376, 'S', '1377-1378, 'S', '1379-1380, 'S', '1381-1382, 'S', '1383-1384, 'S', '1385-1386, 'S', '1387-1388, 'S', '1389-1390, 'S', '1391-1392, 'S', '1393-1394, 'S', '1395-1396, 'S', '1397-1398, 'S', '1399-1400, 'S', '1401-1402, 'S', '1403-1404, 'S', '1405-1406, 'S', '1407-1408, 'S', '1409-1410, 'S', '1411-1412, 'S', '1413-1414, 'S', '1415-1416, 'S', '1417-1418, 'S', '1419-1420, 'S', '1421-1422, 'S', '1423-1424, 'S', '1425-1426, 'S', '1427-1428, 'S', '1429-1430, 'S', '1431-1432, 'S', '1433-1434, 'S', '1435-1436, 'S', '1437-1438, 'S', '1439-1440, 'S', '1441-1442, 'S', '1443-1444, 'S', '1445-1446, 'S', '1447-1448, 'S', '1449-1450, 'S', '1451-1452, 'S', '1453-1454, 'S', '1455-1456, 'S', '1457-1458, 'S', '1459-1460, 'S', '1461-1462, 'S', '1463-1464, 'S', '1465-1466, 'S', '1467-1468, 'S', '1469-1470, 'S', '1471-1472, 'S', '1473-1474, 'S', '1475-1476, 'S', '1477-1478, 'S', '1479-1480, 'S', '1481-1482, 'S', '1483-1484, 'S', '1485-1486, 'S', '1487-1488, 'S', '1489-1490, 'S', '1491-1492, 'S', '1493-1494, 'S', '1495-1496, 'S', '1497-1498, 'S', '1499-1500, 'S', '1501-1502, 'S', '1503-1504, 'S', '1505-1506, 'S', '1507-1508, 'S', '1509-1510, 'S', '1511-1512, 'S', '1513-1514, 'S', '1515-1516, 'S', '1517-1518, 'S', '1519-1520, 'S', '1521-1522, 'S', '1523-1524, 'S', '1525-1526, 'S', '1527-1528, 'S', '1529-1530, 'S', '1531-1532, 'S', '1533-1534, 'S', '1535-1536, 'S', '1537-1538, 'S', '1539-1540, 'S', '1541-1542, 'S', '1543-1544, 'S', '1545-1546, 'S', '1547-1548, 'S', '1549-1550, 'S', '1551-1552, 'S', '1553-1554, 'S', '1555-1556, 'S', '1557-1558, 'S', '1559-1560, 'S', '1561-1562, 'S', '1563-1564, 'S', '1565-1566, 'S', '1567-1568, 'S', '1569-1570, 'S', '1571-1572, 'S', '1573-1574, 'S', '1575-1576, 'S', '1577-1578, 'S', '1579-1580, 'S', '1581-1582, 'S', '1583-1584, 'S', '1585-1586, 'S', '1587-1588, 'S', '1589-1590, 'S', '1591-1592, 'S', '1593-1594, 'S', '1595-1596, 'S', '1597-1598, 'S', '1599-1600, 'S', '1601-1602, 'S', '1603-1604, 'S', '1605-1606, 'S', '1607-1608, 'S', '1609-1610, 'S', '1611-1612, 'S', '1613-1614, 'S', '1615-1616, 'S', '1617-1618, 'S', '1619-1620, 'S', '1621-1622, 'S', '1623-1624, 'S', '1625-1626, 'S', '1627-1628, 'S', '1629-1630, 'S', '1631-1632, 'S', '1633-1634, 'S', '1635-1636, 'S', '1637-1638, 'S', '1639-1640, 'S', '1641-1642, 'S', '1643-1644, 'S', '1645-1646, 'S', '1647-1648, 'S', '1649-1650, 'S', '1651-1652, 'S', '1653-1654, 'S', '1655-1656, 'S', '1657-1658, 'S', '1659-1660, 'S', '1661-1662, 'S', '1663-1664, 'S', '1665-1666, 'S', '1667-1668, 'S', '1669-1670, 'S', '1671-1672, 'S', '1673-1674, 'S', '1675-1676, 'S', '1677-1678, 'S', '1679-1680, 'S', '1681-1682, 'S', '1683-1684, 'S', '1685-1686, 'S', '1687-1688, 'S', '1689-1690, 'S', '1691-1692, 'S', '1693-1694, 'S', '1695-1696, 'S', '1697-1698, 'S', '1699-1700, 'S', '1701-1702, 'S', '1703-1704, 'S', '1705-1706, 'S', '1707-1708, 'S', '1709-1710, 'S', '1711-1712, 'S', '1713-1714, 'S', '1715-1716, 'S', '1717-1718, 'S', '1719-1720, 'S', '1721-1722, 'S', '1723-1724, 'S', '1725-1726, 'S', '1727-1728, 'S', '1729-1730, 'S', '1731-1732, 'S', '1733-1734, 'S', '1735-1736, 'S', '1737-1738, 'S', '1739-1740, 'S', '1741-1742, 'S', '1743-1744, 'S', '1745-1746, 'S', '1747-1748, 'S', '1749-1750, 'S', '1751-1752, 'S', '1753-1754, 'S', '1755-1756, 'S', '1757-1758, 'S', '1759-1760, 'S', '1761-1762, 'S', '1763-1764, 'S', '1765-1766, 'S', '1767-1768, 'S', '1769-1770, 'S', '1771-1772, 'S', '1773-1774, 'S', '1775-1776, 'S', '1777-1778, 'S', '1779-1780, 'S', '1781-1782, 'S', '1783-1784, 'S', '1785-1786, 'S', '1787-1788, 'S', '1789-1790, 'S', '1791-1792, 'S', '1793-1794, 'S', '1795-1796, 'S', '1797-1798, 'S', '1799-1800, 'S', '1801-1802, 'S', '1803-1804, 'S', '1805-1806, 'S', '1807-1808, 'S', '1809-1810, 'S', '1811-1812, 'S', '1813-1814, 'S', '1815-1816, 'S', '1817-1818, 'S', '1819-1820, 'S', '1821-1822, 'S', '1823-1824, 'S', '1825-1826, 'S', '1827-1828, 'S', '1829-1830, 'S', '1831-1832, 'S', '1833-1834, 'S', '1835-1836, 'S', '1837-1838, 'S', '1839-1840, 'S', '1841-1842, 'S', '1843-1844, 'S', '1845-1846, 'S', '1847-1848, 'S', '1849-1850, 'S', '1851-1852, 'S', '1853-1854, 'S', '1855-1856, 'S', '1857-1858, 'S', '1859-1860, 'S', '1861-1862, 'S', '1863-1864, 'S', '1865-1866, 'S', '1867-1868, 'S', '1869-1870, 'S', '1871-1872, 'S', '1873-1874, 'S', '1875-1876, 'S', '1877-1878, 'S', '1879-1880, 'S', '1881-1882, 'S', '1883-1884, 'S', '1885-1886, 'S', '1887-1888, 'S', '1889-1890, 'S', '1891-1892, 'S', '1893-1894, 'S', '1895-1896, 'S', '1897-1898, 'S', '1899-1900, 'S', '1901-1902, 'S', '1903-1904, 'S', '1905-1906, 'S', '1907-1908, 'S', '1909-1910, 'S', '1911-1912, 'S', '1913-1914, 'S', '1915-1916, 'S', '1917-1918, 'S', '1919-1920, 'S', '1921-1922, 'S', '1923-1924, 'S', '1925-1926, 'S', '1927-1928, 'S', '1929-1930, 'S', '1931-1932, 'S', '1933-1934, 'S', '1935-1936, 'S', '1937-1938, 'S', '1939-1940, 'S', '1941-1942, 'S', '1943-1944, 'S', '1945-1946, 'S', '1947-1948, 'S', '1949-1950, 'S', '1951-1952, 'S', '1953-1954, 'S', '1955-1956, 'S', '1957-1958, 'S', '1959-1960, 'S', '1961-1962, 'S', '1963-1964, 'S', '1965-1966, 'S', '1967-1968, 'S', '1969-1970, 'S', '1971-1972, 'S', '1973-1974, 'S', '1975-1976, 'S', '1977-1978, 'S', '1979-1980, 'S', '1981-1982, 'S', '1983-1984, 'S', '1985-1986, 'S', '1987-1988, 'S', '1989-1990, 'S', '1991-1992, 'S', '1993-1994, 'S', '1995-1996, 'S', '1997-1998, 'S', '1999-2000, 'S', '2001-2002, 'S', '2003-2004, 'S', '2005-2006, 'S', '2007-2008, 'S', '2009-2010, 'S', '2011-2012, 'S', '2013-2014, 'S', '2015-2016, 'S', '2017-2018, 'S', '2019-2020, 'S', '2021-2022, 'S', '2023-2024, 'S', '2025-2026, 'S', '2027-2028, 'S', '2029-2030, 'S', '2031-2032, 'S', '2033-2034, 'S', '2035-2036, 'S', '2037-2038, 'S', '2039-2040, 'S', '2041-2042, 'S', '2043-2044, 'S', '2045-2046, 'S', '2047-2048, 'S', '2049-2050, 'S', '2051-2052, 'S', '2053-2054, 'S', '2055-2056, 'S', '2057-2058, 'S', '2059-2060, 'S', '2061-2062, 'S', '2063-2064, 'S', '2065-2066, 'S', '2067-2068, 'S', '2069-2070, 'S', '2071-2072, 'S', '2073-2074, 'S', '2075-2076, 'S', '2077-2078, 'S', '2079-2080, 'S', '2081-2082, 'S', '2083-2084, 'S', '2085-2086, 'S', '2087-2088, 'S', '2089-2090, 'S', '2091-2092, 'S', '2093-2094, 'S', '2095-2096, 'S', '2097-2098, 'S', '2099-2100, 'S', '2101-2102, 'S', '2103-2104, 'S', '2105-2106, 'S', '2107-2108, 'S', '2109-2110, 'S', '2111-2112, 'S', '2113-2114, 'S', '2115-2116, 'S', '2117-2118, 'S', '2119-2120, 'S', '2121-2122, 'S', '2123-2124, 'S', '2125-2126, 'S', '2127-2128, 'S', '2129-2130, 'S', '2131-2132, 'S', '2133-2134, 'S', '2135-2136, 'S', '2137-2138, 'S', '2139-2140, 'S', '2141-2142, 'S', '2143-2144, 'S', '2145-2146, 'S', '2147-2148, 'S', '2149-2150, 'S', '2151-2152, 'S', '2153-2154, 'S', '2155-2156, 'S', '2157-2158, 'S', '2159-2160, 'S', '2161-2162, 'S', '2163-2164, 'S', '2165-2166, 'S', '2167-2168, 'S', '2169-2170, 'S', '2171-2172, 'S', '2173-2174, 'S', '2175-2176, 'S', '2177-2178, 'S', '2179-2180, 'S', '2181-2182, 'S', '2183-2184, 'S', '2185-2186, 'S', '2187-2188, 'S', '2189-2190, 'S', '2191-2192, 'S', '2193-2194, 'S', '2195-2196, 'S', '2197-2198, 'S', '2199-2200, 'S', '2201-2202, 'S', '2203-2204, 'S', '2205-2206, 'S', '2207-2208, 'S', '2209-2210, 'S', '2211-2212, 'S', '2213-2214, 'S', '2215-2216, 'S', '2217-2218, 'S', '2219-2220, 'S', '2221-2222, 'S', '2223-2224, 'S', '2225-2226, 'S', '2227-2228, 'S', '2229-2230, 'S', '2231-2232, 'S', '2233-2234, 'S', '2235-2236, 'S', '2237-2238, 'S', '2239-2240, 'S', '2241-2242, 'S', '2243-2244, 'S', '2245-2246, 'S', '2247-2248, 'S', '2249-2250, 'S', '2251-2252, 'S', '2253-2254, 'S', '2255-2256, 'S', '2257-2258, 'S', '2259-2260, 'S', '2261-2262, 'S', '2263-2264, 'S', '2265-2266, 'S', '2267-2268, 'S', '2269-2270, 'S', '2271-2272, 'S', '2273-2274, 'S', '2275-2276, 'S', '2277-2278, 'S', '2279-2280, 'S', '2281-2282, 'S', '2283-2284, 'S', '2285-2286, 'S', '2287-2288, 'S', '2289-2290, 'S', '2291-2292, 'S', '2293-2294, 'S', '2295-2296, 'S', '2297-2298, 'S', '2299-2300, 'S', '2301-2302, 'S', '2303-2304, 'S', '2305-2306, 'S', '2307-2308, 'S', '2309-2310, 'S', '2311-2312, 'S', '2313-2314, 'S', '2315-2316, 'S', '2317-2318, 'S', '2319-2320, 'S', '2321-2322, 'S', '2323-2324, 'S', '2325-2326, 'S', '2327-2328, 'S', '2329-2330, 'S', '2331-2332, 'S', '2333-2334, 'S', '2335-2336, 'S', '2337-2338, 'S', '2339-2340, 'S', '2341-2342, 'S', '2343-2344, 'S', '2345-2346, 'S', '2347-2348, 'S', '2349-2350, 'S', '2351-2352, 'S', '2353-2354, 'S', '2355-2356, 'S', '2357-2358, 'S', '2359-2360, 'S', '2361-2362, 'S', '2363-2364, 'S', '2365-2366, 'S', '2367-2368, 'S', '2369-2370, 'S', '2371-2372, 'S', '2373-2374, 'S', '2375-2376, 'S', '2377-2378, 'S', '2379-2380, 'S', '2381-2382, 'S', '2383-2384, 'S', '2385-2386, 'S', '2387-2388, 'S', '2389-2390, 'S',

```
Db      375  LeuAsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheaspGly 394
QY      1249 GAGCTTCAGAGGGTCACGTGCTTGTTGGTGGCCACAGAACTTCGCCCTCGCTGGAAAGTCC 1308
Db      395  GluLeuGlnGluGlyValLeuIleValProGlnAsnPheAlaValAlaIalysSer 414
QY      1309 CAGAGCGAGACTTCAATACGTGGCATTCAGACAGACTCAAGGCCACGATAGCCAC 1368
Db      415  GlnSerAspAsnPhelutyrValSerPheLysThrAsnAspArgProSerIleGlyAsn 434
QY      1369 CTCGCCGTGAAACCTCCGTCTATAGATAACCTCCGAGGAGGTGTTGCAAAATTCATAT 1428
Db      435  LeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnHisThrPhe 454
QY      1429 GGCCTTCAAAGGAGCAGGACGAGCTTAAGAACAAACACCCCTTCAAGTTCTTCGTT 1488
Db      455  AsnLeuLysSerGlnGlnAlaArgGlnValLysAsnAsnAsnPropheSerPheLeuVal 474
QY      1489 CCACCGTCTCAGCAGTCTCCGAGGCGTGGCT 1521
Db      475  ProProGlnGluSerGlnArgAlaValAla 485

RESULT 5
FW5YG2
glycinin chain AlaBx precursor - soybean
N;Alternate names: 11S globulin; glycinin AlaB1b
C;Species: Glycine max (soybean)
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 05-Oct-2004
C;Accession: A23497; S10502
R;Negoro, T.; Momma, T.; Fukazawa, C.
Nucleic Acids Res. 13, 6719-6731, 1985
A;Title: A cDNA clone encoding a glycinin A-1a subunit precursor of soybean.
A;Reference number: A23497; MUID:86041867; PMID:2997720
A;Accession: A23497
A;Molecule type: mRNA
A;Residues: 1-495 <NEG>
A;Cross-references: UNIPROT:P04776; UNIPARC:UPI0000042425; GB:X02985; NID:G18614; PIDN:G
A;Experimental source: cv. Boninmori
A;Note: The authors translated the codon AAC for residue 449 as Lys
R;Kitamura, Y.; Arahira, M.; Itoh, Y.; Fukazawa, C.
Nucleic Acids Res. 18, 4245, 1990
A;Title: The complete nucleotide sequence of soybean glycinin A2B1a gene spanning to and
A;Reference number: S10502; MUID:90332420; PMID:2377465
A;Accession: S10502
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 481-495 <KIT>
A;Cross-references: UNIPARC:UPI00000067E; EMBL:X53404; NID:G18522; PIDN:CAA37479.1; PID
C;Comment: The source of this protein was cotyledon tissue taken 38 days after flowering
C;Keyword: seed; storage protein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-306/Product: glycinin chain Ala #status predicted <GLA>
F;311-490/Product: glycinin chain Bx #status predicted <GLB>
F;107-317/Disulfide bonds: #status predicted

Alignment Scores:
Pred. No.: 6,32e-101 Length: 495
Score: 1466.00 Matches: 288
Percent Similarity: 70.1% Conservative: 73
Best Local Similarity: 55.9% Mismatches: 104
Query Match: 52.8% Indels: 50
DB: 1 Gaps: 6

US-10-728-323-3 (1-1524) x FWSYG2 (1-495)

QY      1  CGGCACACCGGAGGAGACCGTCCAGTTCAGCGCTCAATGGCAGAGACCTGAC 60
Db      23  ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProGly 42
QY      61  AATCGCATTTGAATCAGAGGCGGTTCATTGAGACTTGGAACTTGGACCCCAACACACGAGGAGTTC 120
```

```
Db      43  AsnArgIleGluSerGluGlyGlyLeuIleGluThrTrpAsnProAsnAsnLysProPhe 62
QY      121  GAATGCGCGCGTCCCTCTCTCGCTTAGTCTCTCCGCGCAACGCCCTTCGTAGGCCT 180
Db      63  GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 82
QY      181  TTCTACTCAAAAGCTCCCGCAGGAGATCTTCATCCAGCAGGAGGGGATACTTTGGTTG 240
Db      83  SerTyrThrAsnGlyProGlnGluIleIleGlnGlnGlyLysGlyIlePheGlyMec 102
QY      241  ATATTCCTCTGTTGCTTCTAGACACTATGAAGAGCCCTCACACAAGGTCGTGCATCTCAG 300
Db      103  IleTyrProGlyCysSerSerThrPheGluGluProGlnGlnProGlnGlnArgGlyGln 122
QY      301  TCCAAAAGACCAACAGAGCTCTCAAGGAGAAGCAACCAAGCAACAGCAGACAGATAGT 360
Db      123  SerSerArgPro-----GlnAspArg 129
QY      361  CACCAAGAGGTGCACGGTTTCGATGAGGTGATCTCATTTGCAGTTCCCAACCGGTGTGCT 420
Db      130  HisGlnLysIleTyrAsnSerArgGluGlyAspLeuIleAlaValProThrGlyValAla 149
QY      421  TTCTGCTCTACACGACCCACGACACTGATGTTGTTGTTCTTCTTCTACTGACCAAC 480
Db      150  TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerIleIleAspThrAsn 169
QY      481  AACACGACCAACAGCTTGATCAGTCCCGCAGAGATTCAATTTGGCTGGGACACGGAG 540
Db      170  SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 189
QY      541  CAAGAGTTCTTAAGGTACCAGCAACAAAGCAGACAAAGCAGCAGCAAGAAAGCTTACCATAT 600
Db      190  GlnGluPheLeuLysTyrGlnGlnGln----- 199
QY      601  AGCCCATACAGCCCGCAAAAGTCAGCCTAGACAAGAGCGCTGAATTTAGCCCTCGAGGA 660
Db      200  -----GlyGly 201
QY      661  CAGCACAGCGCAGAGAACGACGAGGACAAGAGAAGAAAAGAGGTGGAACATCTTC 720
Db      202  HisGlnSerGlnLysGlyLysHisGlnGlnGluGluAsnGluGlySerIleLeu 221
QY      721  ACGCGCTTCACGCGGAGTCTCTGGAAACAAGCTCTCCAGTTGACGACACAGACATAGTG 780
Db      222  SerGlyPheThrLeuGluPheLeuGluHisAlaPheSerVal---AspLysGlnIleAla 240
QY      781  CAAACCTAAGAGCGGACCGAGAGTGAAGAGAGGAGGCCATTGTGACGTAGGGGA 840
Db      241  LysAsnLeuGlnGlyGluAsnGluGlyGluAspLysGlyAlaIleValThrValLysGly 260
QY      841  GGCCTCAGAACTTTGAGCCCA-----GATAGAAGAGAGCTCCGACGAGAAGAGAG 891
Db      261  GlyLeuSerValIleLysProThrAspGluGlnGlnArgProGlnGluGlu 280
QY      892  GAATACGATCAAGATGAATATGAATAC-----GATGAAGAG---GATAGAAGG 936
Db      281  GluGluGluLysGluLysProGlnCysLysGlyLysAspLysHisCysGlnArgPro 300
QY      937  CGTGCGAGGGAGCAGAGCGAGGGGAATGTGATTGAAGAGACGATCTGCACCGCAAGT 996
Db      301  ArgGlySerGlnSerLysSerArgArgAsnGlyIleAspGluThrIleCysThrMetArg 320
QY      997  GCTAAAAGAACATTGGTAGAAACAGATCCCTGACATCTACACCTCAAGCTGTTCA 1056
Db      321  LeuArgHisAsnIleGlyGlnThrSerSerProAspIleTyrAsnProGlnAlaGlySer 340
QY      1057  CTCAAACTGCCAACCGATCTCAACCTTCTAATACTTAGTGGCTTGGACCTAGTGTGAA 1116
Db      341  ValThrThrAlaThrSerLeuAspPheProAlaLeuSerTrpLeuArgLeuSerAlaGly 360
QY      1117  TATGGAATCTCTACAGGAATGCAATTGTTGTCGCTCACTACAAACCAACGACACAGC 1176
Db      361  PheGlySerLeuArgLysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSer 380
```


QY 1381 AACTCCGTATAGATAACCTGCGGAGGAGGTGGTTGCAAAATTCATATGCGCTCCAAAGG 1440
Db SerSerValIleAspAspLeuProLeuAspValValAlaAlaThrPheAsnMetGlnArg 471
QY 1441 GAGCAGCGCAGGAGGAGTAAAGAACAAACACCCCTTCAAGTGTCTTCGCTCCACCGTCTCAG 1500
Db AsnGluAlaArgGlnLeuLysSerAsnAsnProPheLysPheLeuValProProArgGln 491

QY 1501 CAGTCTCCGAGGCGTGTGGCT 1521

Db 492 SerGluMetArgAlaSerAla 498

RESULT 7

S08237

legumin A2 precursor - garden pea

N;Alternate names: 11S seed storage protein

C;Species: Pisum sativum (garden pea)

C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 05-Oct-2004

C;Accession: S08237

R;Revie, W.G.; Whitecross, M.I.; Higgins, T.J.V.

Nucleic Acids Res. 18, 655, 1990

A;Title: Nucleotide sequence of an A-type legumin gene from pea.

A;Reference number: S08237; MUID:90174993; PMID:2308850

A;Accession: S08237

A;Molecule type: DNA

A;Residues: 1-520 <RER>

A;Cross-references: UNIPROT:P15838; UNIPARC:UPI000012E424; EMBL:X17193; NID:g20773; PIDN

C;Genetics:

A;Gene: legA2

A;Introns: 97/1; 180/3; 391/3

F;1-22/Domain: signal sequence #status predicted <SIG>

F;23-335/Product: legumin A2 alpha chain #status predicted <ACH>

F;336-520/Product: legumin A2 beta chain #status predicted <BCH>

Alignment Scores:

Pred. No.:	4,17e-100	Length:	520
Score:	1455.00	Matches:	295
Percent Similarity:	65.5%	Conservative:	64
Best Local Similarity:	53.8%	Mismatches:	97
Query Match:	52.4%	Indels:	92
DB:	2	Gaps:	8

US-10-728-323-3 (1-1524) x S08237 (1-520)

QY 1 CGCAGCAACCGGAGGAGAACCGGTCCAGTTCAGCGCTCAATGCGAGACCTGAC 60
Db 24 ArgGluGlnProGluGlnAsnGluCysGlnLeuGluArgLeuAsnAlaLeuGluProAsp 43
QY 61 AATCGATTGAATCAGAGGCGGTACATTGAGACTTGGNACCCCAACACAGGAGTTC 120
Db 44 AsnArgIleGluSerGluGlyGlyLeuIleGluThrTrpAsnProAsnAsnLysGlnPhe 63
QY 121 GAATGCGCGCGCTCGCCCTCTCTCGTCTAGTCTCCGCGCAACGCCCTTCGTAGCGCT 180
Db 64 ArgCysAlaGlyValAlaLeuSerArgAlaThrLeuGlnHisAsnAlaLeuArgPro 83
QY 181 TTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGAGTACTTTGGTTG 240
Db 84 Tyr-Tyr-SerAsnAlaProGlnGluIlePheIleGlnGlyAsnGlyTyrPheGlyMet 103
QY 241 ATATTCCGTGGTGTCTCTAGACACTATGAGAGCGCTCACACAAAGTCTGCTCATCTCAG 300
Db 104 ValPheProGlyCysProGluThrPheGluGluPro-----GlnGlu 117
QY 301 TCCCAAGACCAACCAAGACGCTCCAAAGGAGAGAGACCAAGCAACAGCAACGAGATAGT 360
Db 118 SerGlu-----GlnGlyGlu-----GlyArgArgTyrArgAspArg 129
QY 361 CACAGAGGTGCACCGTTTCGATGAGGGTGATCTCATTCGAGTTCACCGGTGTGCT 420
Db 130 HisGlnLysValAsnArgPheArgGluGlyAspIleAlaValProThrGlyIleVal 149

QY 421 TTCTGCTCTCAACAGCACGACACATGATGTTGTTGTTCTTCTTCTTACTGACACCAAC 480
Db 150 PheTrpMetCysAsnAspGlnAspThrProValIleAlaValSerLeuThrAspIleArg 169
QY 481 AACAACGACAACAGCGTTCAGTTCCTCCAGGAGATTCAATTTGGCTGGGACACGGAG 540
Db 170 SerSerAsnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnHisGlu 189
QY 541 CAAAGTCTTTAAGTACCAGCAACAAAGCAGACAAAGCAGACGAGAAAGCTTACCATAT 600
Db 190 GlnGluPheLeuArgTyrGlnHisGln----- 198
QY 601 AGCCCATACAGCCCGCAAAAGTCAGCCTAGACAAAGAGCGTCAATTTAGCCCTCGAGGA 660
Db 198 ----- 198
QY 661 CAGCACAGCGCAGAGACGAGCAGGACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 199 -----GlnGlyGlyLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 213
QY 721 AGCGGCTTCACGCGGAGTTCCTGGAACAAGCCTTCAGGTTGACGACACAGACAGATAGTG 780
Db 214 SerGlyPheLysArgAspPheLeuGluAspAlaPheAsnVal---AsnArgHisIleVal 232
QY 781 CAAAACCTAAGAGCGGAGACCGAGAGTGAAGAGAGGAGGAGCCATTCTGACAGTGAGGGGA 840
Db 233 AspArgLeuGlnGlyArgAsnGluAspGluGluLysGlyAlaIleValLysValLysGly 252
QY 841 GCGCTCAGAATCTTGAGCCCGACATAGAAAG----- 870
Db 253 GlyLeuSerIleIleSerProGluLysGlnAlaArgHisGlnArgGlySerArgGln 272
QY 871 -----AGACGTCCGACGAA 885
Db 273 GluGluAspGluAspGluAspGluArgGlnProArgHisGlnArgGlySerArgGln 292
QY 886 GAAGAGGAATACGATGAAGATGAATAT-----GAA 915
Db 293 GluGluGluGluAspGluAspGluArgGlnProArgHisGlnArgArgArgGlyGlu 312
QY 916 TAGCATGAAGAGGATAGAGCGTGGCAGGGGAAGC-----AGAGGC 957
Db 313 GluGluGluGluAspLysLysGluArgArgGlySerGlnLysGlyLysSerArgArgGln 332
QY 958 AGGGGAATGTTATCAAGAGACGATCTGCACCGCAAGTCTATAAAGACATTCGTAGA 1017
Db 333 GlyAspAsnGlyLeuGluGluThrValCysThrAlaLysLeuArgLeuAsnIleGlyPro 352
QY 1018 AACAGATCCCTGACATCTACAACCTCAAGCTGGTTCCTCAAACTGCCAACGATCTC 1077
Db 353 SerSerSerProAspIleTyrAsnProGluAlaGlyArgIleLysThrValThrSerLeu 372
QY 1078 AACCTTCTAATCTTAGTGGCTTGGACCTAGTGTGAATATATGAAATCTCTACAGAAT 1137
Db 373 AspLeuProValLeuArgTrpLeuLysLeuSerAlaGluHisGlySerLeuHisLysAsn 392
QY 1138 GAATTTGTCGCTCCTACACCAACGACGACATCATATATCGATTGAGGGGA 1197
Db 393 AlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeuLysGly 412
QY 1198 CGGGCTCAGTGCACACTCGTGACAGCAACGCGGACAGAGTGTACACGAGGAGCTCAA 1257
Db 413 ArgAlaArgLeuGlnValValAsnCysAsnGlyAsnThrValPheAspGlyGluLeuGlu 432
QY 1258 GAGGTCACGTGCTTGTGGTGCACAACTTCGCGTCCGTGCGTGGAAAGTCCAGAGCGAG 1317
Db 433 AlaGlyArgAlaLeuThrValProGlnAsnTyrAlaValAlaAlaLysSerLeuSerAsp 452
QY 1318 AACTTCGATACGTGCGATTCCAGACAGACTCAAGGCCCGACGATCCCACTCCGCGGT 1377
Db 453 ArgPheSerTyrValAlaPheLysThrAsnAspArgAlaGlyIleAlaArgLeuAlaGly 472
QY 1378 GAAAACTCCGTCATAGATAACCTCGCGGAGGAGGTGTTGCAAAATTCATATGCGCTCCAA 1437


```
Db 472 SerValIleAsnAsnLeuProLeuAspValAlaAlaThrPheAsnLeuGlnArgAsn 491
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1444 CAGGCAAGGAGGCTTAAGAACAAACACCCCTTCAAGTCTTCTCCACCGTCTCAGCAG 1503
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 492 GluAlaArgGlnLeuLysSerAsnAsnProPheLysPheLeuValProAlaArgGluSer 511
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1504 TCTCCGAGGCTGTGGCT 1521
||||| |||
Db 512 GluAsnArgAlaSerAla 517
||||| |||

RESULT 9
S49877
legumin A precursor - Vicia narbonensis
C:Species: Vicia narbonensis
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 05-Oct-2004
C:Accession: S49877
R:Nong, V.; Becker, C.; Muentz, K.
submitted to the EMBL Data Library, November 1994
A:Description: PCR cloning of legumin cDNA from Vicia narbonensis.
A:Reference number: S49877
A:Accession: S49877
A:Molecule type: mRNA
A:Residues: 1-482 <NON>
A:Cross-references: UNIPROT:Q41676; UNIPARC:UPI000000A6C09; EMBL:Z46803; NID:g600107; PIR
A:Experimental source: tissue type cotyledon
C:Keywords: seed; storage protein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-482/Product: legumin A #status predicted <MAT>

Alignment Scores:
Pred. No.: 2,3e-98 Length: 482
Score: 1431.50 Matches: 284
Percent Similarity: 68.2% Conservative: 66
Best Local Similarity: 55.4% Mismatches: 104
Query Match: 51.5% Indels: 59
DB: 2 Gaps: 7

US-10-728-323-3 (1-1524) x S49877 (1-482)
QY 1 CGGCAGCAACCGGAGGAAACGGCTGCCAGTTCAGCGCTCAATCGCAGAGACCTGAC 60
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 23 ArgGluGlnSerGlnGlnAsnGluCysGlnLeuGluArgLeuAspAlaLeuGluProAsp 42
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 61 AATCGCATTAATCAGAGGCGGTACATTGAGACTTGGAACCCCAACCAACAGGAGTTC 120
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 43 AsnArgIleGluSerGluGlyGlyLeuIleGluThrTriPAsnProAsnAsnArgGlnPhe 62
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 121 GAATGCGCGGCTCGCCCTCTCTCGCTTAGTCTCGCGCAACGCCCTTCGTAGGCCT 180
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 63 ArgCysAlaGlyValAlaLeuSerArgValThrLeuGlnArgAsnAlaLeuArgArgPro 82
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 181 TTCTACTCCAAATGCTCCCAAGAGATCTTCATCCAGCAAGGAAGGGGATACTTTGGGTG 240
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 83 TyrTyrSerAsnAlaProGlnGluIleTyrIleGlnGlnGlyAsnGlyTyrPheGlyVal 102
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 241 ATATTCCTGGTGTCTCTAGACACTATGAAGAGCCTCACACAAAGTCTGCTCATCTCAG 300
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 103 ValPheProGlyCysProGluThrPheGluGluPro-----GlnGluSer 117
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 301 TCCCAAGACCAACCAAGACCTCTCCAAGGAGAGAACCAAGCAACAGCAACAGAGATAGT 360
||||| |||||
Db 118 GluGlnArgGluArgArg------TyrArgAspSer 128
||||| |||||
QY 361 CACCAGAAGGTGACCCCTTTTCGATGAGGTGATCTCATTCAGTTCCTCCACCGGTGTGCT 420
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 129 HisGlnLysValAsnArgPheArgGluGlyAspIleAlaValProThrGlyAsnVal 148
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 421 TTCTGGCTCTAACACGACACGACACTGATGTGTGCTCTTCTCTTACTGACCAAC 480
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 149 LeuTrpMetTyrAsnAspGlnAspThrProValIleAlaIleSerLeuThrAspThrGly 168
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 481 AACCAACGACCAACAGCTTGATCATGTTCCCCAGGAGATTCATTTGGCTGGGAACACGAG 540
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
```

RESULT 10
S11003

```
Db 169 SerSerAsnAsnGlnLeuAspGlnIleProArgArgPheTyrLeuAlaGlyAsnGlnGlu 188
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 541 CAAGAGTTCTTTAAGGTACAGCAACAAAGACAGACAAAGACAGACGAAAGAGCTTACCATAT 600
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 189 GlnGluPheLeuArgTyrGln----- 195
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 601 AGCCCATACAGCCCGCAAAAGTCTAGCTAGACAAGAAGCGGTGAATTTAGCCCTCGAGGA 660
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 195 ----- 195
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 661 CAGCACAGCCCGCAGAGACGAGCAGGA---CAAGAAGAAGAAAAAGAGGTGGAACATC 717
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 196 -----ArgGluGlnGlyGlyGlnGlnGlnGlnGlnAsnAspGlyAsnAsnIle 211
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 718 TTCAGCGGCTTCACGCCGAGGATTCCTGGAAACAAGCTTCCAGGTTCACGACAGACAGATA 777
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 212 PheSerGlyPheLysArgAspPheLeuGluAspAlaLeuAsnVal---AsnArgHisIle 230
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 778 GTGCAAAACCTAAGAGGCGCAGACCGAGTGAAGAGAGGGAGCCATTGTGACAGTGAGG 837
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 231 ValAspArgLeuGlnGlyArgAsnGluAspGluGlyAlaIleValLysValLys 250
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 838 CGAGGCTCAGAAATCTTGAGCCCATAGAAAGACGCTGCCGACGAAAGAGGAATAC 897
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 251 GlyGlyLeuSerIleIleThrProGluArgGlnArgGlySerArgGlnGluGlu--- 269
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 898 GATGAAGATGAATATGATATACGATGAAGAGGATAGAAGCGCTGGCGGGAAGACAGAGGC 957
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 270 AspGluAspGluLysGluGlnProSerArgArgAspGluSerGlnLysGly 289
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 958 AGGGG-----AATGTTATTGAAGAGACCATCTGCACCCCAAGTGCTAAA 1002
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 290 GluSerArgArgHisGlyAspAsnGlyLeuGluGluThrValCysThrAlaLysLeuArg 309
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1003 AAGAACATTGGTAGAAACAGATCCCTGACATCTCAACACCTCAAGCTGGTTCACCTCAA 1062
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 310 ValAsnIleGlySerSerProSerProAspIleTyrAsnProGlnAlaGlyArgIleAsn 329
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1063 ACTGCCAAGCATCTCAACCTTCTAATACCTTAGTGGCTGGACCTAGTCTGAATATGGA 1122
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 330 ThrValThrSerLeuAspLeuProValLeuArgTrpLeuLysLeuSerAlaGluHisGly 349
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1123 AATCTCTACAGGAATGCTATTGTCGCTCACTCAACACCAACCAACGACACAGCATCATA 1182
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 350 SerLeuArgLysAsnAlaLeuIleValProHisTyrAsnArgAsnAlaAsnSerValIle 369
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1183 TATCGATTGAGGCGCGGCTCACGTGCAAGTCTGCGACGCAACCGCAACGAGGTGATC 1242
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 370 TyrAlaLeuLysGlyArgAlaArgLeuGlnValValAsnCysAsnGlyAsnThrValPhe 389
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1243 GACGAGGAGCTTCAAGAGGTCACGTGCTTGTGTCGCCACAGAACTTCGCGCTCGTGGGA 1302
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 390 AspGlyGluLeuAlaGlyArgAlaLeuThrValProGlnAsnTyrAlaValAlaAla 409
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1303 AAGTCCACAGGCGAGAACTTCGAATACGTGGCATTCAGACAGACTCAAGGCCACAGATA 1362
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 410 LysSerLeuSerGluArgPheThrTyrValAlaPheLysThrAsnAspArgAspGlyIle 429
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1363 GCCAACTCCCGCGTGAAACTCCGTCATAGATAACTCCCGGAGGAGGTGGTTGCAAAAT 1422
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 430 AlaArgLeuAlaGlyThrSerValIleAsnAspLeuProLeuAspValValAlaAla 449
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1423 TCATATGGCTCTCAAGGAGGAGCGGACAGGACCTTAAGAACCAACACCCCTCAAGTTC 1482
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 450 ThrPheAsnLeuGlnArgAsnGluAlaArgGlnLeuLysSerAsnAsnProPheLysLeu 469
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1483 TTCGTTCCACCGCTCTCAGCAGTCTCCGAGGCGTGTGCGCT 1521
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 470 LeuValProArgGluSerGluLysArgAlaSerAla 482
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
```

glycinin G3 precursor - soybean
 C:Species: Glycine max (soybean)
 C:Date: 21-Nov-1993 #sequence_revision 19-Jan-1996 #text_change 05-Oct-2004
 C:Accession: S11003
 R:Nielsen, N.C.; Dickinson, C.D.; Cho, T.J.; Thanh, V.H.; Scallan, B.J.; Fischer, R.L.;
 Plant Cell 1, 313-328, 1989
 A:Title: Characterization of the glycinin gene family in soybean.
 A:Reference number: S10851; MUID:92393391; PMID:2485233
 A:Accession: S11003
 A:Molecule type: DNA
 A:Residues: 1-484 <NIE>
 A:Cross-references: UNIPARC:UPI0000177DFE
 A:Experimental source: variety Dare
 C:Genetics:
 A:Gene: GY3
 C:Keywords: storage protein
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-484/Product: glycinin G3 #status predicted <NAT>

Alignment Scores: 4.18e-98 Length: 484
 Pred. No.: 1428.00 Matches: 290
 Score: 70.3% Conservative: 68
 Percent Similarity: 57.0% Mismatches: 95
 Best Local Similarity: 51.4% Indels: 56
 Query Match: 2 Gaps: 9
 DB:

US-10-728-323-3 (1-1524) x S11003 (1-484)

QY	1	CGGAGCAACCGAGAGAACGCGTGCAGCTCCAGCGCTCAATGCGGAGAGACTGAC	60
Db	23	ArgGluGlnProGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp	42
QY	61	AATCGATTGAATCAGAGCGGCTTACATTGAGCTTGAAACCCCAACACAGAGTTC	120
Db	43	AsnArgileGluSerGluGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe	62
QY	121	GAATGCGCGCGGCTCCCTCTCTCGCTTAGTCTCCGCGCAACGCGCTTCGTAGGCT	180
Db	63	GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro	82
QY	181	TTCTACTCAATGCTCCCGAGAGATCTTCATCCAGAAAGGAGATCTTTGGTTG	240
Db	83	SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet	102
QY	241	ATATTCCCTGGTTGCTTAGACTATGAGAGCCTCACACAAGGTCGTGATCTCAG	300
Db	103	IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln	119
QY	301	TCCCAAGACCAACCAAGAGCTCTCCAAGGAGAAGACCAAGCCACAGCAACGAGTAGT	360
Db	120	SerSerArgPro-----GlnAspArg	126
QY	361	CACCAAGAGGTGCACGGTTTCGATGAGGTGATCTCATTCAGTTCACCGGTGTGCT	420
Db	127	HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla	146
QY	421	TTCTGGCTCTACACACACGACACTGATGTTGTGCTTCTTCTTACTGACACCAAC	480
Db	147	TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn	166
QY	481	NACAACGACACACAGCTTCAGTTCCCGAGGATTCATATTGGCTGGGAACCGAG	540
Db	167	SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu	186
QY	541	CAAGAGTTCTTAAGGTACCAACAAAGCAGCAACAAAGCAGACGAAAGACTTACCATAT	600
Db	187	GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly-----	199
QY	601	AGCCCATACAGCCCGCAAAAGTCAGCCTAGACAAAGAGCGGTGAATTTAGCCCTCGAGA	660
Db	200	-----Gly	200

QY	661	CAGCACAGCCGAGAGAAACGAGCAGGACAAAGAAGAAAGAAACGAAAGTGGAAACATCTTC	720
Db	201	ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluAsnGluGlySerIleLeu	220
QY	721	ACGCGCTTCACGCGGAGTTCTCGAACAAAGCTTCAGGTTGACGACAGACAGATAGTG	780
Db	221	SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal	239
QY	781	CAAAACCTTAAGAGGCGAGCAGCAGAGTAGTGAAGAAGAGGAGGAGCATTGTGACAGTGAAGGA	840
Db	240	ArgLysLeuGlnGlyGluAsnGluGluGluGlyAlaIleValThrValLysGly	259
QY	841	GGCTCAGAAATCTTGAGCCCA-----GATAGAAAGAGACGTGCGCAGCAAGAAGAG	891
Db	260	GlyLeuSerValIleSerProThrGluGlnGlnGlnArgProGluGluGlu	279
QY	892	GAATACGATGAAGATGAATATGATATACATGATGAAGAGGATAGAGCGTGGCAGGGGAGC	951
Db	280	LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer	294
QY	952	AGAGCAGGGGGAATGTTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAGAAACATC	1011
Db	295	Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle	311
QY	1012	GGTAGAAACAGATCCCTGACATCTACAAACCTCAAGCTGCTCACTCAAACTGCCAAC	1071
Db	312	GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr	331
QY	1072	GATCTCAACCTTCTAATACTTGTAGTGGCTTGACCTAGTCTGATATATGGAATCTCTAC	1131
Db	332	SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg	351
QY	1132	AGGAATGATTGTTTGTCTGCTCCTACACCAACCAACGACACAGCATCATATATCGATTG	1191
Db	352	LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu	371
QY	1192	AGGGACCGGCTCAGCTGCAAGTCTGACACAGCAACGCGCAACAGAGTGTACGACGAGGAG	1251
Db	372	AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu	391
QY	1252	CTTCAAGAGGTGACGCTGCTGTGTGTCACAGAACTTCGCGCTCGCTGGAAAGTCCGAG	1311
Db	392	LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln	411
QY	1312	AGCGAAGACTTCGAATAGTGTGCTCAAGACAGACTCAAGSCCCAGCATAGCCCAACCTC	1371
Db	412	SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu	431
QY	1372	GCCGTGAAAACTCCGTCATAGATAACCTCGCGGAGGAGGTGGTTGCAAAATTCATATGCG	1431
Db	432	AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn	451
QY	1432	CTCCAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1488
Db	452	LeuArgGlnSerGlnValSerGluLeuLysTyrGluGlyAsnTrpGlyProLeuValAsn	471
QY	1489	CCACGCTCTCAGCAG---TCTCCGAGG	1512
Db	472	ProGluSerGlnGlnGlySerProArg	480

RESULT 11
 S14393
 legumin A2 precursor - fava bean
 C:Species: Vicia faba (fava bean)
 C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
 C:Accession: S14393
 R:Schlesier, B.; Bassuener, R.; van Hal, N.; Muentz, K.
 Nucleic Acids Res. 18, 7146, 1990
 A:Title: The cDNA derived primary structure of two distinct legumin A subunit precursors:
 A:Reference number: S14392; MUID:91088307; PMID:2263481
 A:Accession: S14393
 A:Status: preliminary
 A:Molecule type: mRNA

A;Residues: 1-500 <SCH>
A;Cross-references: UNIPROT:Q99304; UNIPARC:UPI00000A2504; EMBL:X55014; NID:g22007; PIDN
C;Superfamily: glycinin

Alignment Scores: 4.22e-97 Length: 500
Pred. No.: 1414.50 Matches: 280
Score: 67.3% Conservative: 76
Percent Similarity: 52.9% Mismatches: 100
Best Local Similarity: 50.9% Indels: 73
Query Match: 2 Gaps: 8
DB:

US-10-728-323-3 (1-1524) x S14393 (1-500)

```
QY 1 CGGCAGACCGGAGGAGAACGGCTGCCAGTTCACGCGCTCAATCGCGAGACCTGAC 60
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
23 ArgGluGlnSerGlnGlnAsnGluCysGlnLeuGluArgLeuAspAlaLeuGluProAsp 42
QY 61 AATCGCATTCATCAGAGCGGCTTACATTGAGACTTGAACCCCAACACAGGAGTTC 120
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
43 AsnArgIleGluSerGluGlyGlyLeuIleGluThrTrpAsnProAsnAsnArgGlnPhe 62
QY 121 GAATGCGCGCGCTGCCCTCTCTCGCTTAGTCTCCGCGCAGCAACGCCCTTCGTAGCCT 180
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
63 ArgCysAlaSerValAlaLeuSerArgAlaThrLeuGlnArgAsnAlaLeuArgArgPro 82
QY 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGATACTTTGGTTG 240
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
83 TyrTy-SerAsnAlaProGlnGluIleTyrIleGlnGlnGlyAsnGlyTyrPheGlyMet 102
QY 241 ATATTCCCTGGTTGCTCTAGACACTATGAGAGCCTTCACACAAAGTTCGTCATCTAG 300
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
103 ValPheProSerCysProGluThrPheGluGluPro-----GlnGln 116
QY 301 TCCCAAGACCAACCAAGACGCTCTCCAGGAGAGAGACCAAGCCCAACAGCAACGAGATAGT 360
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
117 SerGlu-----GlnGlyGlu-----GlyGlyArgTyrArgAspSer 128
QY 361 CACCAGAGGTGCACCGTTTCGATGAGGGTGTCTCATTCAGATTCACCGCGGTGTGCT 420
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
129 HisGlnLysValAsnArgPheArgGluGlyAspIleAlaValProThrGlyIleVal 148
QY 421 TTCTGCTCTACACACACGACACTGTATGTTGCTGCTTCTCTCTCTCTCTCTCTCTCT 480
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
149 PheTrpMetTyrAsnAspGlnAspThrProValIleAlaIleSerLeuThrAspIleGly 168
QY 481 AACAAACGACAAACAGCTTCATTCAGTTCCTCCAGGAGATTCAAATTTGGCTGGGAACACGGAG 540
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
169 SerSerAsnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 188
QY 541 CAAGAGTTCCTTAAGGTACAGCAACAAAGCAGACAAAGCAGCAAGAGCTTACCATAT 600
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
189 GlnGluPheLeuArgTyrGlnHisGln-----GlnGlnGlnGlnGlnGlnGlnGln 197
QY 601 AGCCCATACAGCCCGCAAGTACAGCTAGCAGACAGAGAGCGTGAATTAGCCCTCGGGA 660
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
197 -----GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 197
QY 661 CAGCAGACCGCGCAGAGAACGAGCAGGACAGAGAAAGAAAGAGGTGGAAACATCTTC 720
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
198 -----GlnGlyGlyLysGluGluGlnAspAsnAspGlyAsnAsnIlePhe 212
QY 721 AGCGGTTTACGCGGAGTTCCTGGAAACAAGCTTCCAGGTTGACGACAGAGATGTG 780
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
213 SerGlyPheLysArgAspPheLeuGluAspAlaLeuAsnVal---AsnArgHisIleVal 231
QY 781 CAAACCTAGAGCGGACCGAGAGTGAAGAGAGGAGCCCATTTGACAGTACGGGGA 840
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
232 AspArgLeuGlnGlyArgAsnGluAspGluGluLysGlyAlaIleValLysValLysGly 251
QY 841 GGCCTCAGAAATCTTGAGCCCGCAGATAGAAAG-----AGACGTGCGCGAGAA 885
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
252 GlyLeuSerIleIleThrProGluArgGlnAlaArgHisProArgGlySerArgGln 271
```

```
QY 886 GAAGAGGAATACGATCAAGATCAATATGAA----- 915
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
272 GluGluAspGluAspGluAspGluGluArgGlnProSerHisLysSerArg 291
QY 916 -----TACGATGAAGAGGATAGA-----AGCGTGGCAGGGGAAACGAGA 954
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
292 ArgGlyGluAspGluAspAspLysGluLysArgHisSerGlnLysGlyGluSerArg 311
QY 955 GGCAGGGGAATGGTATTGAAGAGACGATCTCCACCGCAAGTGTCTAAAGAACATTTGGT 1014
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
312 HisGlyAspAsnGlyLeuGluGluThrValCysThrAlaLysLeuArgLeuAsnIleGly 331
QY 1015 AGAAACAGATCCCTCGACATCTACACCTCAAGCTGCTCACTCAAACTCCCAACGAT 1074
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
332 SerSerSerSerProAspIleTyrAsnProGlnAlaGlyArgIleLysThrValThrSer 351
QY 1075 CTCACCTTCTAATCTAGTGGCTTGACCTAGTCTGTAATATGGAATCTCTACAGG 1134
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
352 LeuAspLeuProValLeuArgTrpLeuLysLeuSerAlaGluHisGlySerLeuArgLys 371
QY 1135 AATGATCTGTTTGTCTCATCAACACCAACGACACAGCATCATATATCGATTGAGG 1194
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
372 AsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleLeuTyrAlaLeuLys 391
QY 1195 GGACGGCTCAGTCCAGTCTGGACACGACCAACGACAGAGTGTACGAGGAGCTT 1254
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
392 GlyArgAlaArgLeuGlnValValAsnCysAsnGlyAsnThrValPheAspGluGluLeu 411
QY 1255 CAAGAGGGTACGCTGCTGTCGACAGAACTTCGCGCTCGCTGGAAGTCCCAAGAGC 1314
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
412 GluAlaGlyArgAlaLeuThrValProGlnAsnTyrAlaValAlaLysSerLeuSer 431
QY 1315 GAGAACTTCGAATACGTGTCATCAAGACAGACTCAAGGCCGACGATACCACTCGCC 1374
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
432 AspArgPheThrTyrValAlaPheLysThrAsnAspArgAlaGlyIleAlaArgLeuAla 451
QY 1375 GGTGAAACTCCGTCATAGATAACCTCGCGAGGAGGTGGTTCGCAATTCATATGCCCTC 1434
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
452 GlyThrSerSerValIleAsnAspMetProValAspValValAlaAlaThrPheAsnLeu 471
QY 1435 CAAAGGAGCAGCAGGAGGAGCTTAAGAAACAACACCCCTTCAAGTTCCTCCACCG 1494
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
472 GluArgAsnGluAlaArgGlnLeuLysSerAsnAsnProPheLysPheLeuValProPro 491
QY 1495 TCTCAGCAGTCTCCGAGGCTGTGGCT 1521
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
492 ArgGluSerGlnLysArgAlaSerAla 500
```

RESULT 12

T08452
probable legumin A precursor - garden pea
C;Species: Pisum sativum (garden pea)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-Oct-2004
C;Accession: T06452
R;Lycett, G.W.; Croy, R.R.D.; Shirsat, A.H.; Richards, D.M.; Boulter, D.
Nucleic Acids Res. 13, 6733-6743, 1985
A;Title: The 5'-flanking regions of three legumin genes: comparison of the DNA sequence
A;Reference number: Z15687; MUID:86041868; PMID:2997721
A;Accession: T06452
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-507 <b1yC>
A;Cross-references: UNIPARC:UPI000016DF29; EMBL:X02982; NID:g20777; PIDN:CAA26720.1; P11
C;Genetics:
A;Intons: 96/1; 179/3; 388/3
C;Keywords: seed; storage protein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-332/Product: legumin A, alpha chain #status predicted <ALP>

Alignment Scores: 7.79e-95 Length: 507
Pred. No.: 1384.00 Matches: 283
Score:

Percent Similarity:	63.9%	Conservative:	66
Best Local Similarity:	51.8%	Mismatches:	97
Query Match:	49.8%	Indels:	100
DB:	2	Gaps:	9
US-10-728-323-3 (1-1524) x T06452 (1-507)			
Qy	1	CGGCACACCGGAGGAGAACCGGTGCGCAGTTCAGCGCCTCAATGCGCAGACCTTGAC	60
Db	23	ArgGluGlnProGlnGlnAsnGluCysGlnLeuGluArgLeuAspAlaLeuGluP:coAsp	42
Qy	61	AATCGCATTGAATCAGAGCGCGGTTCATATTGAGACTTGGACCCCAACACAGGAGTTC	120
Db	43	AsnArgIleGluSerGluGlyLeuIleGluThrTrpAsnProAsnAsnLysGlnPhe	62
Qy	121	GAATGCGCGCGGTGCGCCTCTCTCGCTTAGTCTCCGCGCAACGCGCCTTCGTAGGCCT	180
Db	63	ArgCysAlaGlyValAlaLeuSerArgAlaThrLeuGlnArgAsnAlaLeuArgArgPro	82
Qy	181	TTCTACTCAATGCTCCCGAGAGATCTTCATCCAGCAGAGGAGGGATCTTTGGGTTG	240
Db	83	TyrTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyAsnGlyTyrPheGlyMet	102
Qy	241	ATATTCCCTGGTTGCTTAGACACTATGAAGAGCTCACAACAAGTCTCGATCTCAG	300
Db	103	ValPheProGlyCysProGluThrPheGluGluPro-----GlnGlu	116
Qy	301	TCCCAAGACCCACCAAGCGTCTCCAAGGAGAAGACCAAGCCAAACGACGAGTAGT	360
Db	117	SerGlu-----GlnGlyGlu-----GlyArgArgTyrArgAspArg	128
Qy	361	CACGAGAGGTGCACGGTTTCATGAGGGTGATCTCATTCAGTTCACCGGTGTGCT	420
Db	129	HisGlnLysValAsnArgPheArgGluGlyAspIleAlaValProThrGlyIleVal	148
Qy	421	TTCTGGCTCTACACGACACGACTGATGTTGTTCTGTTCTTCTTACTGACACCAAC	480
Db	149	PheTrpMetTyrAsnAspGlnAspThrProValIleAlaValSerLeuThrAspIleArg	168
Qy	481	AACACGACCAACACGAGTTCAGTTCCTCCAGGAGATTCATTTGGTGGGAACACGGAG	540
Db	169	SerSerAsnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnHisGlu	188
Qy	541	CAAGAGTCTTAAAGGTACCAGCAACAAAGACAGACAAAGACGACGAGAGCTTACCATAT	600
Db	189	GlnGluPheLeuGlnTyrGlnHisGln-----	197
Qy	601	AGCCCATACAGCCCGCAAGTCAGCTCCTAGACACAGAGCGTGAATTTAGCCCTCAGGA	660
Db	197	-----	197
Qy	661	CAGCACAGCCGACAGACGACGAGCAAGAAAGAAACGAAGTGGAAACATCTTC	720
Db	198	-----GlnGlyClyLysGlnGlnGlnAsnGluGlyAsnAsnIlePhe	212
Qy	721	AGCGGCTTCACGCGGAGTTCCTGGAACAACGCTTCAGGTGTGACGACAGACATAGTG	780
Db	213	SerGlyPheLysArgAspTyrLeuGluAspAlaPheAsnVal---AsnArgHisIleVal	231
Qy	781	CMAAECTAAGAGGCGAGACCAGAGTGAAGAAGAGGAGCCATTGTGACAGTGAAGGGA	840
Db	232	AspArgLeuGlnGlyArgAsnGluAspGluGlyGlyAlaIleValLysValLysGly	251
Qy	841	GGCCTCAGATCTTCAGGCCCATAGAAG-----	870
Db	252	GlyLeuSerIleIleSerProGluLysGlnAlaArgHisGlnArgGlySerArgGln	271
Qy	871	-----AGACGTGCGCAGCAAGAAAGAG	891
Db	272	GluGluAspGluAspGluLysGlnProArgHisGlnArgGlySerArgGlnGlu	291
Qy	892	GAATACGATGAAGATGAATATGATACGATGAAGAGGATAGAAGCGGTGGC-----	942

Db 1 CysGlnLeuGluArgLeuAspAlaLeuGluProAsnArgIleGluSerGluGlyGly 20
QY 85 TACATTGAGACTTGGAAACCCCAACACAGAGAGTTCGAATTCGCGCGCGCTCGCCCTCTCT 144
Db 21 LeuIleGluThrTrpAsnProAsnAsnArgGlnPheArgCysAlaArgValAlaLeuSer 40
QY 145 CGTTTAGTCTCCGCGCGCAACGCGCTTCGTAGCCCTTCTATCTCCCAATGCTCCCAAGGAG 204
Db 41 ArgAlaThrLeuGlnArgAsnAlaLeuArgArgProTyrTyrSerAsnAlaProGlnGlu 60
QY 205 ATCTTCATCCAGCAAGGAGGATCTTGGGTTCATATCCCTGGTGTCTCTAGACAC 264
Db 61 IleTyrIleGlnGlnGlyAsnGlyTyrPheGlyMetValPheProSerCysProGluThr 80
QY 265 TATGAAGAGCCTCACACAAAGTCTGCTCATCTCAGTCCCAAGACACCAAGACGCTCTC 324
Db 81 PheGluGluPro-----GlnGlnSerGlu----- 88
QY 325 CAAGGAGAGACCAAGCCAAACAGACGAGTAGTACACAGAGTGTGACCGTTCGAT 384
Db 89 GlnGlyGlu-----GlyArgArgTyrArgAspSerHisGlnLysValAsnArgPheArg 106
QY 385 GAGGGTATCTCATTCGAGTCCACCGGTGTGCTTCTCGCTCTACACGACACACGAC 444
Db 107 GlnGlyAspIleIleAlaValProThrGlyIleValPheTrpMetTyrAsnAspGlnAsp 126
QY 445 ACTGATGTGTGCTGCTTCTTCTTACTGACCAACAAACACGACCAACGAGCTTGATCAG 504
Db 127 IleProValIleAlaIleSerLeuThrAspThrGlySerSerAsnAsnGlnLeuAspGln 146
QY 505 TTCCCGCAGGAGATTCAATTGGCTGGGAACACGAGCAAGAGTCTTAAAGTACCAAGCAA 564
Db 147 MetProArgArgPheTyrLeuAlaGlyAsnGlnGluGlnPheLeuArgTyrGlnHis 166
QY 565 CAAAGCAGACAAAGCAGACGAAGAAGCTTACCATATAGCCCATACACCGCGCAAGTCCAG 624
Db 167 GlnGlnGlyValLysGluGlu-----GlnAspAsn 176
QY 625 CCTAGACAAGAGCGTGAATTTAGCCCTCGAGGACGACACGCGCGCAGAGAACGAGCA 684
Db 177 AspGlyAsnGlnGlnGluPheLeu---ArgTyrGlnHis-----ArgGlnGlyVal 193
QY 685 GGACAGAGAAGAAACGAGAGTGGAAACATCTTCAGCGGCTTCAGCGCGGAGTCTCTG 744
Db 194 LysGluGlnGlnAspAsnAspGlyAsnAsnIlePheSerGlyPheAsnArgAspPheLeu 213
QY 745 GAACAAGCCTTCAGGTTGACGACACACATAGTCAAAACCTTAAGAGCGGAGCCGAG 804
Db 214 GluAspAlaPheAsnVal---AsnArgHisIleValAspArgLeuGlnGly-----Arg 230
QY 805 AGTGAAGAAGAGGAGCCATTGTGACAGTGGAGGGGCGCTCAGAAATCTTGAGCCCGAGAT 864
Db 231 AsnGluGluArgGlyAlaIleValLysValLysGlyGlyLeuSerIleIleThrProPro 250
QY 865 AGAAGAGA-----CGTGCCGACGAAAGAGGAATACGATGAA 903
Db 251 GluArgGlnAlaArgHisProArgLysArgGlySerArgGlnGluGluAspGluAspGlu 270
QY 904 GATTGATATGAA-----TACGATGAAGAG 927
Db 271 AspGluLysGluGluArgGlnProSerHisIleLysSerArgArgGlyGluAspGluAsp 290
QY 928 GATAGA-----AGCGCTGGCAGGGGAAGCAGACGCGCGGGAATGTGATT 972
Db 291 AspLysGluLysArgHisSerGlnLysGlyGluSerArgArgHisGlyAspAsnGlyLeu 310
QY 973 GAAGACGATCTGCACCGCAAGTCTAAAGAACATTCGTAGAAACACATCCCTGAC 1032
Db 311 GluGluThrValCysThrAlaLysLeuArgLeuAsnIleGlySerSerSerProAsp 330
QY 1033 ATCTACAACCTCAAGCTGGTTCACATCAAACTGCCCAACCATCTCAACCTTCTTAATACTT 1092

Db 331 IleTyrAsnProGlnAlaGlyArgIleLysThrValThrSerValAspLeuProValLeu 350
QY 1093 AGGTGGCTTGGACCTAGTGTCTGAATATGGAATCTCTACAGGAATGCATTGTTTTCGCT 1152
Db 351 ArgTyrPLeuLysLeuSerAlaGluHisGlySerLeuArgLysAsnAlaMetPheValPro 370
QY 1153 CACTACACCAACGACGACACATCATATATCGATTGAGGGGACGGCTCAGCTGCAA 1212
Db 371 HisTyrAsnLeuAsnAlaAsnSerValLeuTyrAlaLeuLysGlyArgAlaArgLeuGln 390
QY 1213 GTCTGGACAGCAACGGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTCACTGCTT 1272
Db 391 ValValAsnCysAsnGlyAsnThrValPheAspGlyGluLeuGluAlaGlyArgAlaLeu 410
QY 1273 GTGTGTCACAGAACTTCGCCCTCGTGGAAAGTCCAGGACGAGCAACTTCGAATACGTG 1332
Db 411 ThrValProGlnAsnTyrValValAlaAlaLysSerLeuSerAspArgPheThrTyrVal 430
QY 1333 GCATTCACAGACACTCAAGGCCAGCATAGCAACCTCGCGGTGAAACTCCGTCATA 1392
Db 431 AlaPheLysThrAsnAspArgAlaGlyIleAlaIleArgLeuAlaGlyThrSerSerValIle 450
QY 1393 GATAACCTCGCGAGAGGTGTTGCAAAATTCATATGGCTCCAAAGGGAGGAGCAAG 1452
Db 451 AsnAspLeuProLeuAspValValAlaAlaThrPheAsnLeuGluArgAsnGluAlaArg 470
QY 1453 CAGCTTAAGAACAAACCCCTTCAGTTCCTTCGTTCCACCGCTCTCAG 1500
Db 471 GlnLeuLysPheAsnAsnProSerArgPheLeuValProArgGlu 486
RESULT 14
S51941
prunin 1 precursor - almond
N;Alternate names: salt-soluble globulin; seed storage protein
C;Species: Prunus dulcis (almond)
C;Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 05-Oct-2004
C;Accession: S51941; S42474
R;Garcia-Mas, J.; Messesguer, R.; Arus, P.; Puigdomenech, P.
Plant Mol. Biol. 27, 205-210, 1995
A;Title: Molecular characterization of cDNAs corresponding to genes expressed during al
A;Reference number: S51940; MUID:95170003; PMID:7865791
A;Accession: S51941
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-551 <GAR>
A;Cross-references: UNIPROT:Q43607; UNIPARC:UPI0000177DF8; EMBL:X78119
A;Note: the source is designated as Prunus amygdalus
R;Garcia-Mas, J.; Messesguer, R.; Arus, P.; Puigdomenech, P.
submitted to the EMBL Data Library, March 1994
A;Description: Molecular characterization of cDNAs corresponding to proteins related to
A;Reference number: S42473
A;Accession: S42474
A;Molecule type: mRNA
A;Residues: 1-60, 'G', '62-551 <GAW>
A;Cross-references: UNIPARC:UPI00000ACDD6; EMBL:X78119; NID:g460805; PIDN:CAA55009.1; P
C;Keywords: seed; storage protein
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-367/Product: prunin 1 alpha chain #status predicted <MAT1>
F;368-551/Product: prunin 1 beta chain #status predicted <MAT2>
F;108-374/Disulfide bonds: #status predicted
Alignment Scores:
Pred. No.: 9.02e-72 Length: 551
Score: 1073.50 Matches: 228
Percent Similarity: 57.2% Conservative: 89
Best Local Similarity: 41.2% Mismatches: 160
Query Match: 38.6% Indels: 77
DB: 2 Gaps: 11
US-10-728-323-3 (1-1524) x S51941 (1-551)
QY 4 CAGCAACCGGAGGAGAACGCGTGCCAGTTCACGCGCTCAATGCGCAGACCTGACAAT 63


```
QY 79 GCGGTTCATTGAGCTGGACCCCAACACACAGGAGTTCGAATGCGCGCGCTCGCC 138
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
50 GlyGlyLeuIleGlnThrTrpAsnSerGlnHisProGluLeuLysCysAlaGlyValThr 69
QY 139 CTCCTCGCTTAGCTCCCGCCCAACAGCCCTTCGTAGGCTTCTTACTCCCAATGCTCCC 198
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
70 ValSerLysLeuThrLeuAsnArgAsnGlyLeuHisLeuProSerTyrSerProTyrPro 89
QY 199 CAGGAGATCTTCATCAGCAAGGAGGATATCTTTGGTGTGATATCTCCCTGGTGTCTCT 258
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
90 ArgMetIleIleAlaGlnGlyLysGlyAlaLeuGlyValAlaIleProGlyCysPro 109
QY 259 AGACACTATCAAGAGCTCACACACAGGTGCTGATCTCAGTCCCAAGACACCAAGA 318
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
110 GluThrPheGluGluProGlnGluGlnSerAsnArgGlySerArg----- 125
QY 319 CGTCTCCAAGGAGAACCAAGCCCAACAGACAGATAGTCACCAAGAGGTGCAACCGT 378
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
126 -----SerGlnLysGlnGlnLeuGlnAspSerHisGlnLysIleArgHis 140
QY 379 TTCATGAGGTGATCTCATTTGAGTTCACCGGTGTTGCTTCTTGGCTCTACACAGAC 438
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
141 PheAsnGluGlyAspValLeuIleProProGlyValProTyrTrpThrTyrAsnThr 160
QY 439 CACGACACTGATGTTCTGCTGTTCTCTTACTGACACCAACACACACAGCAGCTT 498
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
161 GlyAspGluProValValAlaIleSerLeuLeuAspThrSerAsnPheAsnAsnGlnLeu 180
QY 499 GATCAGTTCCTCCAGGAGATTCATTTGGCTGGGAACACGAGCAAGAGTTCCTTAAGGTAC 558
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 AspGlnThrProArgValPheTyrLeuAlaGlyAsnProAspIleGluTyr----- 197
QY 559 CAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATATAGCCCATACGCCGCAA 618
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
198 -----ProGluThrMetGln 202
QY 619 AGTCAGCCTAGACAAGACGCTGAATTTAGCCCTCGAGGACAGACAGCCCGCAGAA 678
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
203 GlnGlnGlnGlnGlnLysSerHisGlyArgLysGlnGlyGlnHisGln----- 219
QY 679 CGAGCAGGACAAGAAGAAACGAAAGGTGGAAACATCTTCACGCGCTTCACGCGGAG 738
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
220 -----GlnGluGluGluGluGlySerValLeuSerGlyPheSerLysHis 236
QY 739 TTCCTGAAACAGCCTTCAGGTTCGACAGACAGATAGTGCACAACTTAAGCGGAG 798
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
237 PheLeuAlaGlnSerPheAsnThrAsnGlu--AspIleAlaGluLysLeu-----Gln 253
QY 799 ACCGAGAGTCAAGAAGGGAGCCATTGTGACAGTGGGGGAGGCTCAGAACTTTGAGC 858
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
254 SerProAspGluArgLysGlnIleValThrValGluGlyLeuSerValIleSer 273
QY 859 CCATAGAGAAGAGACTGCGCAGAGAGGAAGGAATACCATGAA----- 903
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
274 ProLysTrpGlnGlnGlnAspGluAspGluAspGluAspGluAspGluAspGlu 293
QY 904 -----GATGAATAT 912
Db GlnIleProSerHisProProArgProSerHisGlyLysArgGluGlnAspGluAsp 313
QY 913 GAATACGATGAAGGATAGA--AGCGTGGGAGGGAGGAGCAGAGC----- 957
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
314 GluAspGluAspGluAspLysProArgProSerArgProSerGlnGlyLysArgGluGln 333
QY 957 ----- 957
Db 334 AspGlnAspGlnAspGluAspGluAspGluAspGlnProArgLysSerArgGlu 353
QY 958 -----AGGGG 963
Db 354 TrpArgSerLysLysThrGlnProArgProArgProArgGlnGluProArgGluArgGly 373
QY 964 -----AATGTTATTGAAGAGACCATCTGCACCGCAAGTGCTAAAAAGAACATT 1011
```

```
Db 374 CysGluThrArgAsnGlyValGluGluAsnIleCysThrLeuLysLeuHisGluAsnIle 393
QY 1012 GGTAGAAACAGATCCCTCGACATCTACAAACCTCAAGCTGGTGTCTACTCAAAATCGCCAC 1071
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
394 AlaArgProSerArgAlaAspPheTyrAsnProLysAlaGlyArgIleSerThrLeuAsn 413
QY 1072 GATCTCAACCTTCTAATACTAGGTGGCTTGGACCTAGTGTGAATATGAAATCTCTAC 1131
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
414 SerLeuThrLeuProAlaLeuArgGlnPheGlnLeuSerAlaGlnIleTyrValValLeuTyr 433
QY 1132 AGCAATGCATTTGCTGCTCACTACAAACACACGACACACAGCATCATATATCGATTG 1191
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
434 LysAsnGlyIleTyrSerProHisTrpAsnLeuAsnAlaAsnSerValIleTyrValThr 453
QY 1192 AGGGACGGGCTCACGTGCAAGTCTGTGGACAGCAACGCGCAACAGAGTGTACGACGAGGAG 1251
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
454 ArgGlyGlnGlyLysValArgValValAsnCysGlnGlyAsnAlaValPheAspGlyGlu 473
QY 1252 CTTCAAGAGGCTCACGTGCTTGTGTCACAGAACTTCGCGCTCGCTGGAAGTCCCGAG 1311
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
474 LeuArgArgGlyGlnLeuLeuValProGlnAsnPhaValValAlaGluGlnAlaGly 493
QY 1312 AGCGAAGACTTCCAAATACGTGCAATTCAGACAGACTCAAGGCCCGCAGCATAGCAACCTC 1371
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
494 GluGlnGlyPheGluTyrIleValPheLysThrHisHisAsnAlaValThrSerTyrLeu 513
QY 1372 GCGGTGAAACTCCGCTCATAGTAACCTGCGGAGGAGGTGGTTCGCAAAATTCATATGCC 1431
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
514 -----LysAspValPheArgAlaIleProSerGluValLeuAlaHisSerTyrAsn 530
QY 1432 CTCCAAGGGAGCAGCAGGAGGAGCTTAAAGAACACACCCCTTCAAGTTCCTTCGTTCGA 1491
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
531 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 550
QY 1492 CCGTCTCAGCAGCTCTCCGAGGCTGTGGCT 1521
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
551 ProLysGluSerGlnArgValValAla 560
```

Search completed: May 15, 2006, 22:22:27
Job time : 80.0021 secs

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 21:23:21 ; Search time 69.4413 Seconds
(without alignments)
4645.191 Million cell updates/sec

Title: US-10-728-323-3
Perfect score: 2779
Sequence: 1 cggcgacgaacggaggagaa.....ctccgagggtgtggtcttaa 1524

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delopt 6.0 , Delopt 7.0	

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p model -DEV=xlp
-Q-/abses/ABSSWEB_spool/US10728323/runat_15052006.172132.22377/app_query.fasta_1
-DB=uniprot -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss02p
-USER=US10728323 @CGN 1.1.694 @runat 15052006.172132.22377 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOF=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2673	96.2	507	082580	ARAHY
2	2579.5	92.8	531	Q516T2	ARAHY
3	2479	89.2	529	Q9FZ11	ARAHY
4	2455	88.3	536	Q647H4	ARAHY
5	2451.5	88.2	537	Q647H3	ARAHY
6	2444	87.9	538	Q8LKN1	ARAHY
7	2436	87.7	536	Q672T4	ARAHY
8	2421	87.1	530	Q9SQH7	ARAHY
9	1834	66.0	510	Q61WG5	ARAHY
10	1651	59.4	484	Q647H2	ARAHY
11	1502	54.0	481	GLYG3	SOYBN
12	1499.5	54.0	485	GLYG2	SOYBN
13	1499.5	54.0	485	Q549Z4	SOYBN
14	1496	53.8	495	GLYG1	SOYBN
15	1489	53.6	481	Q852U5	SOYBN
16	1488.5	53.6	482	Q852U4	SOYBN

17	1457.5	52.4	498	2	Q41702	VICSA
18	1455	52.4	520	1	LEGA2	PEA
19	1443	51.9	517	2	Q9T0P5	PEA
20	1440	51.8	517	1	LEGA	PEA
21	1431.5	51.5	482	2	Q41676	VICNA
22	1414.5	50.9	500	2	Q99304	VICFA
23	1375	49.5	497	2	Q03971	VICFA
24	1186	42.7	496	2	Q9SMJ4	CICAR
25	1110.5	40.0	515	2	Q8W1C2	CORAV
26	1074.5	38.7	488	2	Q41128	QUERO
27	1069.5	38.5	551	2	Q43607	PRUDU
28	1059	38.1	542	2	Q8LK20	CASCR
29	1049	37.7	560	2	Q9S9D0	SOYBN
30	1034.5	37.2	457	2	Q8GZP6	ANAO
31	1032	37.1	503	1	LEGJ	PEA
32	999.5	36.0	566	2	Q24294	PEA
33	998	35.9	485	2	Q41703	VICSA
34	996.5	35.9	484	1	LEGB4	VICFA
35	986	35.5	517	2	Q39922	GLYSO
36	986	35.5	517	2	Q7GC77	SOYBN
37	985	35.4	517	2	Q9SB12	SOYBN
38	981.5	35.3	564	2	Q43673	VICFA
39	980	35.3	517	2	Q93707	SOYBN
40	980	35.3	517	2	Q93708	SOYBN
41	978.5	35.2	219	2	Q8LLO3	ARAHY
42	975	35.1	563	2	Q9SB11	SOYBN
43	974	35.0	563	2	Q39921	GLYSO
44	973.5	35.0	562	1	GLYG4	SOYBN
45	968	34.8	504	2	Q43608	PRUDU

ALIGNMENTS

RESULT 1
082580 ARAHY PRELIMINARY; PRT; 507 AA.
AC 082580;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Glycinin (Fragment).
GN Name:Arach;
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99146968; PubMed=10021462;
RA Rabin P., Helm E.M., Stanley J.S., West C.M., Sampson H.A.,
RA Burks A.W., Bannon G.A.;
RT "Molecular cloning and epitope analysis of the peanut allergen Ara h 3";
RL J. Clin. Invest. 103:535-542(1999).
DR EMBL; AF093541; AAC63045.1; -, mRNA.
DR HSSP; P04776; IUD1.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin_1; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S_SEED_STORAGE; UNKNOWN_1.
FT NON TER 1
SQ SEQUENCE 507 AA; 58349 MW; F3FB38BC3CB82DED CRC64;
Alignment Scores:
Pred. No.: 3 29e-178 Length: 507
Score: 2673.00 Matches: 507
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0

Query Match: 96.2% Indels: 0
 DB: 2 Gaps: 0
 US-10-728-323-3 (1-1524) x O82580_ARAHY (1-507)

QY 1 CGGCAGCAACCGAGGAGACGGTCCAGTTCACGCGCTCAATCGCAGACGCTGAC 60
 DB 1 ArgGlnGlnProGluGlnAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 20

QY 61 AATTCGATTGAATCAGAGGCGGTACATTGAGACTTGGAAACCCCAACACGAGGAGTTTC 120
 DB 21 AsnArgIleGluSerGluGlyTyrlleGluThrTrpAsnProAsnAsnGlnGluPhe 40

QY 121 GAATGCGCGCGCTCGCCTCTCTCGCTTAGTCTCTCGCGCGCAACGCCCTTCGTAGGCT 180
 DB 41 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 60

QY 181 TTCTACTCCAATGCTCCCGAGGAGATCTCATCCAGCAAGGAGGGGATACTTTGGGTG 240
 DB 61 PheTySerAsnAlaProGlnGlnIlePheIleGlnGlnGlyArgGlyTyrlleGlyLeu 80

QY 241 ATATTCCCTGTTGCTCTAGACACTATGAAGAGCCTCACACAAGGTCTGATCTCAG 300
 DB 81 IlePheProGlyCysProArgHisTyrlleGluProHisThrGlnGlyArgSerGln 100

QY 301 TCCCAAGACCAACAGAGCTCTCCAAGGAGAGACCAAGCCCAACAGCAACGAGATAGT 360
 DB 101 SerGlnArgProArgLeuGlnGlyGluAspGlnSerGlnGlnGlnArgAspSer 120

QY 361 CACGAGAGTGACCGTTTCAGATGAGGTGATCTCATTCAGTTCGCCCGGTGCT 420
 DB 121 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 140

QY 421 TTCTGCTCTACCAACGACCACTGATGTTGTTGCTCTCTCTCTCTCTCTCTCTCTCT 480
 DB 141 PheTrpLeuTyrlleAsnAspHisAspThrAspValAlaValSerLeuThrAspThrAsn 160

QY 481 AACACGACCAACGAGTTCAGTTCCTCCCGAGAGATTCATTTGGCTGGGAACGAG 540
 DB 161 AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnThrGlu 180

QY 541 CAAGAGTCTTAAAGTTACAGCAACAAAGCAGACAAAGCAGACGAGAGAGCTTACCATAT 600
 DB 181 GlnGluPheLeuArgTyrlleGlnGlnSerArgGlnSerArgArgSerLeuProTyrlle 200

QY 601 AGCCCATACAGCCGCAAGTTCAGCTTAGACAAAGAGCGTGAATTTAGCCCTCGAGGA 660
 DB 201 SerProTyrlleSerProGlnSerGlnProArgGlnGluArgGluPheSerProArgGly 220

QY 661 CAGCAGCCGCGAGACGAGCAGGACAAAGAGAGAGAAACCAAGGTGGAACATCTTC 720
 DB 221 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluAsnGluGlyGlyAsnIlePhe 240

QY 721 AGCGGCTTACGCGCGAGTTCTTGGACACAGCTTCAGGTTCCAGCAGACAGACAGATAGT 780
 DB 241 SerGlyPheThrProGluPheLeuGlnGlnAlaPheGlnValAspAspArgGlnIleVal 260

QY 781 CAAACCTAAGAGCGGAGACCGAGAGTGAAGAGAGGAGCCATTGTGACAGTGAAGGGA 840
 DB 261 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 280

QY 841 GGCTCTAGAATCTTGAAGGAGGAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
 DB 281 GlyLeuArgIleLeuSerProAspArgLysArgAlaAspGluGluGluGluTyrlleAsp 300

QY 901 GAAGATGAATATGAATACGATGAAGAGGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAG 960
 DB 301 GluAspGluTyrlleGluTyrlleGluAspArgLysArgGlyArgGlySerArgGlyArg 320

QY 961 GGAATGGTATTGAAGAGAGCATCTGCACCCGAGGTGCTTAAAGAGCAATCTGTGAGAAC 1020
 DB 321 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 340

QY 1021 AGATCCCCTGACATCTACAACCTCAAGCTGGTTCACCTCAAACTGCCAACGATCTCAAC 1080
 DB 341 ArgSerProAspIleTyrlleArgProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 360

QY 1081 CTTCTAAATCTTAGTGGCTTGGACCTAGTGTGAATATGGAATCTCTACAGGAATGCA 1140
 DB 361 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrlleGlyAsnLeuTyrlleArgAsnAla 380

QY 1141 TTGTTTTCGCTCCTACACAACCGACACACAGCATCATATATCGATTGAGGGACGG 1200
 DB 381 LeuPheValAlaHisTyrlleAsnThrAsnAlaHisSerIleIleTyrlleArgLeuArgGlyArg 400

QY 1201 GCTCAGCTCAGTTCGTGACAGCAACCGCAGACAGTGTACGACGAGGAGCTTCAAGAG 1260
 DB 401 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrlleAspGluGluGlnGlu 420

QY 1261 GGTCACTGCTTGTGTGCTCCACAGAACTTTCGCGCTCGCTGGAAAGTCCCAGAGCGAGAAC 1320
 DB 421 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 440

QY 1321 TTTCGAATACGTGCGCATTCGAAGACAGACTCAAGCCGAGCATATAGCCAACTCCCGCGTGAA 1380
 DB 441 PheGluTyrlleValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 460

QY 1381 AACTCGCTCATAGATAACCTGCCGAGGAGGTGGTTCGAAATTCATATGCGCTCCAAAG 1440
 DB 461 AsnSerValIleAspAsnLeuProGluGluValAlaValAlaAsnSerTyrlleGlyLeuGlnArg 480

QY 1441 GAGCAGGCAAGGCGAGCTTAAAGAACCAACCCCTTCAAGTTCCTCGTTCACCGTCTCAG 1500
 DB 481 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProProSerGln 500

QY 1501 CAGTCTCCGAGCGCTGTGCT 1521
 DB 501 GlnSerProArgAlaValAla 507

RESULT 2
 Q516T2_ARAHY PRELIMINARY; PRT; 531 AA.
 AC Q516T2;
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Arachis hly-4.
 OS Arachis hypogaea (Peanut).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
 OC Arachis
 OC NCBI_TaxID=3818;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Yan Y.S., Lin X.D., Zhang Y.S., Wang L., Wu K., Huang S.Z.;
 RT "Isolation of peanut genes encoding arachins and conglutins by
 RT expressed sequence tags";
 RL Plant Sci. 169:439-445 (2005).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Yan Y., Wang L., Huang S.;
 RT "cDNA clone of peanut seed storage protein gene.";
 RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; A1848698; AAW56067.1; -; mRNA.
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR006045; Cupin.
 DR InterPro; IPR007113; Cupin region.
 DR InterPro; IPR006044; Seedstore_11s.
 DR Pfam; PF00190; Cupin_1; 2.
 DR PRINTS; PR00439; 11SGLOBULIN
 DR PROSITE; PS00305; 11S_SEED_STORAGE; UNKNOWN 1.
 SQ SEQUENCE 531 AA; 60736 MW; 674CA60B45077F0D CRC64;

Alignment Scores: 1.14e-171 Length: 531
 Pred. No.:

Score: 2579.50 Matches: 494
 Percent Similarity: 97.6% Conservative: 2
 Best Local Similarity: 97.2% Mismatches: 11
 Query Match: 92.8% Indels: 1
 DB: 2 Gaps: 1

US-10-728-323-3 (1-1524) x Q516T2_ARAHY (1-531)

QY 1 CGCGAGCAACCGGAGAGAACCGTGCAGTTCACAGCGCTCAATGCGCAGAGACTGAC 60
 DB 24 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
 QY 61 AATCGCATTTGAATCAGAGCGCGGTACATTGAGACTTGGAACCCCAACACAGAGTTTC 120
 DB 44 AsnArgIleGluSerGluGlyGlyTyrlleGluThrTriPAsnProAsnAsnGlnGluPhe 63
 QY 121 GAATGCGCGCGTCCGCTCTCGCTTAGTTCCTCCGCGCAACGCCCTTCGTAGGCT 180
 DB 64 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 83
 QY 181 TTCTACTCCAAATGCTCCCGAGAGATCTTCATCCAGAGGAGGAGTACTTTGGGTG 240
 DB 84 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 103
 QY 241 ATATTCCCTGTTGCTTCTAGACACTATGAAGAGCTTCACACAGAGTTCGTCGATCTCAG 300
 DB 104 IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgArgSerGln 123
 QY 301 TCCCAAGACCCAGAGAGCTCTCCAAGGAGAGACCAAGCAACAGACAGATAGT 360
 DB 124 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnGlnAspSer 143
 QY 361 CACCAAGAGTGCACCGTTTCATGAGGTGATCTCATTCAGTTCCTCCACCGGTGTCCT 420
 DB 144 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 163
 QY 421 TTCTGGCTCTAACAGCACACACTGATGTGTGTGCTTCTTCTTACTGACACCAAC 480
 DB 164 PheTyrLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 183
 QY 481 AACACGCAACACAGCTTCATGTTCCCGAGAGATTCATTTGGCTGGGACGAG 540
 DB 184 AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnHisGlu 203
 QY 541 CAAGAGTCTTAAAGTACAGCAACAAAGACAGACAAAGACAGAGAGTTCACATAT 600
 DB 204 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 223
 QY 601 AGCCCATACAGCCCGCAAGTCCAGCTAGACAAAGAGACGCTGAATTTAGCCCTCAGGA 660
 DB 224 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 243
 QY 661 CAGCAGCGCGCAGAGACGACGACGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
 DB 244 GlnHisSerArgArgGluArgAlaGlyGlnGlnGluGluAsnGluGlyGlyAsnIlePhe 263
 QY 721 AGCGGTTCACGCGCGAGTTCCTGGAACAAAGCTTCAGGTTCAGCACACAGATAGT 780
 DB 264 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 283
 QY 781 CAAACCTTAAGAGCGGACCGAGTGAAGAGAGGAGCGCATTTGTGACGTGAGGGA 840
 DB 284 GlnAsnLeuArgGlyGluAsnGluSerGluGluGluGlyAlaIleValThrValArgGly 303
 QY 841 GGCCTCAGATCTTGAGCCAGATAGAAAGAGACGTCGCGAGAGAGAGAGAGAGATACGAT 900
 DB 304 GlyLeuArgIleLeuSerProAspArgLysArgGlyAlaAspGluGluGluGluTyrAsp 323
 QY 901 GAAGATGAATATGATACCATGAAGAGGATAGAAGCGGTGGCAGGAGAGAGAGAGAG 960
 DB 324 GluAspGluTyrGluTyrAspGluGluAspArgArgArgGlyArgGlySerArgGlyArg 343
 QY 961 GGGATGGTATTGAAGAGACGATCTGACCGCAAGTGTCTAAAGAGAACATTTGGTAGAAC 1020

DB 344 GlyAsnGlyIleGluGluThrIleCysThrAlaSerValLysLysAsnIleGlyArgAsn 363
 QY 1021 AGATCCCTGACATCTACAACCTCAAGCTGTTCACCTCAAAACCTGCAACCATCTCAAC 1080
 DB 364 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 383
 QY 1081 CTTCTAATACCTTAGTGCTTGAGCCTAGTCTGAATATGGAATCTCTACAGGAATGCA 1140
 DB 384 LeuLeuIleLeuAsgTyrLeuGlyLeuSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 403
 QY 1141 TTGTTTTCGCTCACTACAACACCAACGACACACATCATATATCGATTGAGGGACGG 1200
 DB 404 LeuPheValProHisTyrAsnThrAsnAlaHisSerIleIleTyrAlaLeuArgGlyArg 423
 QY 1201 GCTCAGCTGCAAGTCTGTCGACAGCAACGCGACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
 DB 424 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 443
 QY 1261 GGTACGCTGCTGTGTCGCCACAGAACTTCGCCGCTCGCTGGAAAGTCCAGAGCGAGAAC 1320
 DB 444 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerAspAsn 463
 QY 1321 TTCGAATACGTGGCATTCGAACACACTCAAGCCCGACATAGCAACCTCCCGCTGAA 1380
 DB 464 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 483
 QY 1381 AACTCCGTCATAGATAAAGCTCCGCGAGGAGGTGTGCAAAATTCATATGGCCCTCCAAAGG 1440
 DB 484 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 503
 QY 1441 GAGCAGGCAAGG---CAGCTTAAAGAACCAACACCCCTTCAAGTTCCTCCACCGCTT 1497
 DB 504 GluGlnAlaArgGlnGlnLeuLysAsnAsnAsnProPheLysPhePheValProProSer 523
 QY 1498 CAGCAGTCTCCGAGGCGTGTGCT 1521
 DB 524 GlnGlnSerProArgAlaValAla 531

RESULT 3
 Q9FZ11_ARAHY PRELIMINARY; PRT: 529 AA.
 AC Q9FZ11; TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 26, Last annotation update)
 DE Gly1.
 GN Name=Gly1;
 OS Arachis hypogaea (Peanut).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
 OC Arachis.
 OX NCBI_taxID=3818;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Jain A.K., Basha S.M.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF125192; AAC01363.1; -, mRNA.
 DR HSP; P04776; IUC.
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR006045; Cupin.
 DR InterPro; IPR007113; Cupin region.
 DR InterPro; IPR006044; Seedstore_11s.
 DR Pfam; PF00190; Cupin_1; 2.
 DR PRINTS; PR00439; 11SGLOBULIN.
 DR PROSITE; PS00305; 11S_SEED_STORAGE; UNKNOWN_1.
 SQ SEQUENCE 529 AA; 60448 MW; BB6F25BC1D6E06A1 CRC64;

Alignment Scores:
 Pred. No.: 1.2e-164 Length: 529
 Score: 2479.00 Matches: 475
 Percent Similarity: 95.3% Conservative: 9

Best Local Similarity: 93.5% Mismatches: 20
 Query Match: 89.2% Indels: 4
 DB: 2 Gaps: 2

US-10-728-323-3 (1-1524) x Q9FZ11_ARAHY (1-529)

```

QY 1 CGGCAGCAACCGGAGAGACGGTCCAGTTCAGCGGCTCAATCGCGAGACCTGAC 60
DB 25 ArgGlnGlnProGlnGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 44
QY 61 AATCGCATTAATCAGAGCGCGGTACATTGAGACTTGGAACCCCAACACAGGAGTTC 120
DB 45 AsnArgLeuGluSerGluGlyTyrlleGluThrTrpAsnProAsnAsnGlnGluPhe 64
QY 121 GAATGCGCGCGCTCCCTCTCTCGCTTAGTCTCCGCGCGAACCCCTTCGTAGGCTT 180
DB 65 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 84
QY 181 TTCCTACTCCAACTGCCAGGAGATCTTCATCCAGCAAGGAGGGATCTTGGGTG 240
DB 85 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 104
QY 241 ATATTCCTCGTTGTCTCTAGACACTATGAAGACCTCACACAAAGTCTGTCATCTCAG 300
DB 105 IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgGln 124
QY 301 TCCCAAGACCAACAGAGCTCTCCAAGGAGAACCAAGC---CAACAGCAACGAGAT 357
DB 125 SerGlnArgAlaProArgPheGluGlyGluAspGlnSerGlnGlnGlnGlnAsp 144
QY 358 AGTCACGAGAGTGCACCGTTTCGATGAGGGTGATCTCATTCAGTCCCAACCGGT 417
DB 145 SerHisGlnLysValArgArgPheAspGluGlyAspLeuIleAlaValProThrGlyVal 164
QY 418 GCTTTCTGGCTTCACCAACACACAGCACTGATGTTGTTGCTGTTCTCTACTGACAC 477
DB 165 AlaLeuTrpMetPheAsnAspHisAspThrAspValValAlaValSerLeuThrAspThr 184
QY 478 AACAAACAGCAACACAGCTTGATCATGTTCCCGAGAGATTCAATTTGGCTGGGACAG 537
DB 185 AsnAsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnHis 204
QY 538 GAGCAGAGTCTTCAAGGTACAGCAACAAGAGCAGCAAGCAGAGAGAGCTTACCA 597
DB 205 GluGlnGluPheLeuArgTyrGlnGln-----GlnSerArgArgSerLeuPro 221
QY 598 TATAGCCCATACAGCCCGCAAGCTCAGCCTAGACAAGAGAGCGTCAATTTAGCCCTCGA 657
DB 222 TyrSerProTyrSerProGlnSerGlnProArgGlnGluArgGluPheSerProArg 241
QY 658 GGCAGCAGACGCGCAGAGAACGAGCAGGACAGAAAGAAAGAAACCAAGGTGAAACATC 717
DB 242 GlyGlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIle 261
QY 718 TTCAGCGGCTTCAGCGGAGTTCCTGGAACAGCTTCAGGTTCAGCAGACAGAGATA 777
DB 262 PheSerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAspAspArgGlnIle 281
QY 778 GTCAAAACCTTAAGAGCGCAGACGAGTGAAGAGAGGAGCGACCTTGTGACAGTGAGG 837
DB 282 ValGlnAsnLeuArgGlyGluAsnGluSerGluGluGluGlyAlaIleValThrValLys 301
QY 838 GAGAGGCTCAGATCTTGAGCCAGATAGAAAAGAGAGCTGCCAGCAAGAGAGGAAATAC 897
DB 302 GlyGlyLeuArgIleLeuSerProAspArgLysArgGlyAlaAspGluGluGluTyr 321
QY 898 GATGAAGATCAATATGAATAGATGAAGAGATAGAAAGGCGTGGCGGGGAACAGAGGC 957
DB 322 AspGluAspGluTyrGluTyrAspGluAspArgArgArgGlyArgGlySerArgGly 341
QY 958 AGGGGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAAAGAACATTGGTAGA 1017
DB 342 ArgGlyAsnGlyIleGluGluThrIleCysThrAlaSerValLysLysAsnIleGlyArg 361

```

```

QY 1018 AACAGATCCCCTGACATCTAGAACCCCTCAAGCTGGTTCACCTCAAAACTGCCCAACGATCTC 1077
DB 362 AsnArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeu 381
QY 1078 AACCTTCTTAATCTAGGTGGCTTGGACCTAGTGTGTAATATGGAATCTCTACAGGAAT 1137
DB 382 AsnLeuLeuIleLeuArgTrpLeuGlyLeuSerAlaGluTyrGlyAsnLeuTyrArgAsn 401
QY 1138 GCATTGTTTGTCTGCTCACTTACAAACCAACGACACAGCATCATATATGATTGAGGGGA 1197
DB 402 AlaLeuPheValProHisTyrAsnThrAsnAlaHisSerIleIleTyrAlaLeuArgGly 421
QY 1198 CCGGCTCAGCTCAGTCTGAGACAGCAACGCAACAGAGTGTACGACGAGAGCTTCAA 1257
DB 422 ArgAlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGln 441
QY 1258 GAGGGTTCAGTCTGTTGTGTGCCACAGAACTTCGCGCTCGCTGGAAGTCCCAGAGCGAG 1317
DB 442 GluGlyHisValLeuValValProGlnAsnPheAlaValAlaGlySerGlnSerAsp 461
QY 1318 AAATTCGAATACGTGGCATTCAAGACAGACTCAAGGCCAGCATAGCCAACTCCCGCT 1377
DB 462 AsnPheGluTyrValAlaPheLysThrAspSerArgProAsnIleAlaAsnPheAlaGly 481
QY 1378 GAAACTCGCTCATAGATAACCTCGCGAGGAGGTGGTTCGCAATTCATATGCGCTCCAA 1437
DB 482 GluAsnSerIleIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuPro 501
QY 1438 AGGAGCAGGAGGAGGAGCGCTTAAGAACAAACACCCCTTCAAGTCTTCTTCCACCGTCT 1497
DB 502 ArgGluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProProSer 521
QY 1498 CAGCAGTCTCCGAGGCTGTGGCT 1521
DB 522 GlnGlnSerLeuArgAlaValAla 529

```

RESULT 4

Q647H4_ARAHY PRELIMINARY; PRT; 536 AA.

AC Q647H4;

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Arachin Ahy-1.

OS Arachis hypogaea (Peanut).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

OC Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;

OC Arachis.

OX NCBI_TaxID=3818;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Yan Y.S., Lin X.D., Zhang Y.S., Wang L., Wu K., Huang S.Z.;

RT "Isolation of peanut genes encoding arachins and conglutins by

RT expressed sequence tags";

RL Plant Sci. 169:439-445(2005).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RA Yan Y., Wang L., Huang S.;

RT "cDNA clone of peanut seed storage protein gene.";

RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY722685; AAU21490.1; -; mRNA.

DR GO; GO:0045735; F:nutrient reservoir activity; IEA.

DR InterPro; IPR007113; Cupin.

DR InterPro; IPR006044; Seedstore_11s.

DR Pfam; PF00190; Cupin_1; 2.

DR PRINTS; PR00439; 11SGLBULIN.

DR PROSITE; PS00305; 11S SEED STORAGE; UNKNOWN_1.

SQ SEQUENCE 536 AA; 61506 MW; 30DBF638719AEB78 CRC64;

Alignment Scores:

Pred. No.: 5,75e-163 Length: 536
Score: 2455.00 Matches: 471
Percent Similarity: 94.0% Conservative: 14
Best Local Similarity: 91.3% Mismatches: 19
Query Match: 88.3% Indels: 12
DB: 2 Gaps: 3

US-10-728-323-3 (1-1524) x Q647H4_ARAHY (1-536)

QY 1 CGGAGCAACCGGAGGAGACCGTCCAGTTCCAGCGCTCAATGCGCAGACCTGAC 60
DB |||||
QY 24 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
DB |||||
QY 61 AATCGCATTGAATCAGAGCGCGTGTACATTGAGACTTGGAAACCCCAACACAGAGGTTTC 120
DB |||||
QY 44 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 63
DB |||||
QY 121 GAATGCGCGCGTCCCTCTCTCGCTTGTAGTCTCCGCGCAACGCCCTTCGTAGGCT 180
DB |||||
QY 64 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 83
DB |||||
QY 181 TTCTACTCCAAATGCTCCCGAGAGATCTTCATCCAGGAGGAGGAGGATCTTGGGTTG 240
DB |||||
QY 84 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 103
DB |||||
QY 241 ATATTCCCTGTTGTCTAGACACTATGAAGAGCCTCACACAAGGTCGTCTGATCTCAG 300
DB |||||
QY 104 IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgArgHisGln 123
DB |||||
QY 301 TCCCAAGACCAACGAGAGCTTCCAGGAGAGAGACCAAGCCACAGCAGAGATAGT 360
DB |||||
QY 124 SerGlnArgProProArgArgPheGlnGlyGlnAspGlnSerGlnGlnGlnGlnAspSer 143
DB |||||
QY 361 CACAGAGGTCACCGTTTCATGAGGTCATCTCATTCAGTTCCTCCAGCGGTGCTGCT 420
DB |||||
QY 144 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 163
DB |||||
QY 421 TTCTGCTCTACAAGCACACACACTGATGTTGTTGTTGTTCTTCTTACTGACACCAAC 480
DB |||||
QY 164 PheTrpMetTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 183
DB |||||
QY 481 AACACGACCAACGAGTTGATCAGTTCCTCCAGGAGATCAATTTGGCTGGGAAACAGGAG 540
DB |||||
QY 184 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHisGlu 203
DB |||||
QY 541 CAAGAGTTCTTAAGTACAGCAACAAACGACAGCAACAGAGAGCTTACCATAT 600
DB |||||
QY 204 GlnGluPheLeuArgTyrGlnGln-----GlnSerArgArgSerLeuProTyr 220
DB |||||
QY 601 AGCCCATACAGCCGCAAGTCCAGCTAGACAAAGAGCGTGAATTTAGCCCTCCAGGA 660
DB |||||
QY 221 SerProTyrSerProGlnThrGlnProLysGlnGluAspArgGluPheSerProArgGly 240
DB |||||
QY 661 CAGCAGCCGCGAGAGACGACAGAGCAAGAAAGAAACGAAAGGTGGAAACATCTTC 720
DB |||||
QY 241 GlnHisGlyArgArgGluArgAlaGlyGlnGlnGlnGlnGlnGlnGlnGlyAsnIlePhe 260
DB |||||
QY 721 AGCGGCTTCACCGCGAGTCTCTGGNACAGGCTTCCAGGTTGACGACAGACAGATAGT 780
DB |||||
QY 261 SerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAspAspArgGlnIleLeu 280
DB |||||
QY 781 CAAACCTAAGAGCGGAGACCGAGTGAAGAGAGGAGCGCCATTTGTACAGTGAAGGGA 840
DB |||||
QY 281 GlnAsnLeuArgGlyGluAsnGluSerAspGluGlnGlyAlaIleValThrValArgGly 300
DB |||||
QY 841 GCGCTCAGAAATTTGAGCCAGATGAAAGAGA-----CGTCCCGCAGCA 885
DB |||||
QY 301 GlyLeuArgIleLeuSerProAspArgIlysArgGlnGlnTyrGluArgProAspGlu 320
DB |||||
QY 886 GAAGAGGATACGATCAGATCAATGATACGATGAGAG-----CATAGA 933
DB |||||
QY 321 GluGluGluTyrAspGluAspGluTyrGluTyrAspGluGluGluArgGlnHisAspArg 340
DB |||||

QY 934 AGCGTGCAGCGGGAAGCAGAGCGGAGGGAATGGTATTGAAGAGACGATCTGCACCGCA 993
DB |||||
QY 341 ArgArgGlyArgGlySerArgGlySerGlyAsnGlyIleGluGluThrIleCysThrAla 360
DB |||||
QY 994 AGTGCTAAAGAACACATTGGTAGAACAGATCCCTGACATCTCAACCCCTCAAGCTCGT 1053
DB |||||
QY 361 SerPheLysLysAsnIleGlyArgAsnArgSerProAspIleTyrAsnProGlnAlaGly 380
DB |||||
QY 1054 TCACCTCAAAACTGCCAACGATCTCAACCTTCTTAATCTTAGTGCTTGGACCTAGTGT 1113
DB |||||
QY 381 SerLeuLysThrAlaAsnGluLeuAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 400
DB |||||
QY 1114 GAATATGAAATCTCTACAGGAATGCAATGTTTGTTCGCTCTACTACAAACCAACGACAC 1173
DB |||||
QY 401 GluTyrGlyAsnLeuTyrArgAsnAlaLeuPheValProHisTyrAsnThrAsnAlaHis 420
DB |||||
QY 1174 AGCATCATATATCGATTGAGGGGCGGCTCACGTCGAAGTCGTCGAGCAGCAACGCAAC 1233
DB |||||
QY 421 SerIleIleTyrAlaLeuArgGlyArgAlaHisValGlnValValAspSerAsnGlyAsp 440
DB |||||
QY 1234 AGAGTGTACGACGAGGAGCTTCAAGAGGCTCACGTCGTTGTTGGTCCACAGAACTTCGCC 1293
DB |||||
QY 441 ArgValPheAspGluGluGlnGluGlyHisValLeuValValProGlnAsnPheAla 460
DB |||||
QY 1294 GTGCTGGAAGTCCCGAGAGCAATTCGAATACGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1353
DB |||||
QY 461 ValAlaGlyLysSerGlnSerGluAsnPheGluTyrValAlaPheLysThrAspSerArg 480
DB |||||
QY 1354 CCCAGCATAGCAACCTCCGCGTGAACCTCCGTCATAGATTAACCTCCGAGGAGGTTG 1413
DB |||||
QY 481 ProSerIleAlaAsnLeuAlaGlyGluAsnSerPheIleAspAsnLeuProGluGluVal 500
DB |||||
QY 1414 GTTGCAAAATTCATATGGCTCCAAAGGAGGAGCAGCAAGCGAGCTTAAGAACCAACACCCC 1473
DB |||||
QY 501 ValAlaAsnSerTyrGlyLeuProArgGluGlnAlaArgGlnLeuLysAsnAsnAsnPro 520
DB |||||
QY 1474 TTCAGTTCCTTCGTTCCACGCTTCACGAGTCTCCGAGGCTGTGGCT 1521
DB |||||
QY 521 PheLysPhePheValProProSerGluGlnSerLeuArgAlaValAla 536
DB |||||

RESULT 5

Q647H3_ARAHY PRELIMINARY; PRT; 537 AA.

AC Q647H3; (TRENBLrel. 28, Created)

DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)

DE Arachis Any-2.

OS Arachis hypogaea (Peanut).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;

OC Arachis.

OX NCBI_TaxID=3818;

RP NUCLEOTIDE SEQUENCE.

RA Van Y.S., Lin X.D., Zhang Y.S., Wang L., Wu K., Huang S.Z.;

RT "Isolation of peanut genes encoding arachins and conglutins by

RT expressed sequence tags.";

RL Plant Sci. 169:439-445(2005).

RN [1]

RN NUCLEOTIDE SEQUENCE.

RA Van Y., Wang L., Huang S.;

RT "cDNA clone of peanut seed storage protein gene.";

RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY22686; AAU21491.1; -, mRNA.

DR GO; GO:0045735; F:nutrient reservoir activity; IEA.

DR InterPro; IPR006045; Cupin.

DR InterPro; IPR007113; Cupin region.

DR InterPro; IPR006044; Seedstore_11s.

DR Pfam; PF00190; Cupin_1; 2.

DR PRINTS; PR00439; 11SGLOBULIN.

DR PROSITE; PS00305; 11s_SEED_STORAGE; UNKNOWN_1.

SQ SEQUENCE 537 AA; 61532 MW; 3BDD034DEA159657 CRC64;

Alignment Scores:

Pred. No.: 1.01e-162 Length: 537
Score: 2451.50 Matches: 474
Percent Similarity: 93.6% Conservative: 10
Best Local Similarity: 91.7% Mismatches: 20
Query Match: 88.2% Indels: 13
DB: 2 Gaps: 4

US-10-728-323-3 (1-1524) x Q647H3_ARAHY (1-537)

QY 1 CGGCACACCGGAGAGACGGCTGCCAGTCCAGCGCTCAATCGCAGACCTGAC 60
DB 24 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
QY 61 AATCGCATTAATCAGAGCGCGGTACATTGACACTTGGAAACCCCAACACGAGGAGTTC 120
DB 44 AsnArgLeuGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnPhe 63
QY 121 GAATGCGCGCGCTCGCCTCTCTCGTTAGTCTCCGCGCGAACCGCCTTCGTAGGCTT 180
DB 64 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 83
QY 181 TTCCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGAGTACTTTGGGTG 240
DB 84 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlyArgGlyTyrPheGlyLeu 103
QY 241 ATATTCCTCGTGTCTCTAGACACTATGAAGAGCCTCACACAAAGGTCTGTCATCTCAG 300
DB 104 IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgHisGln 123
QY 301 TCCCAAGACCCACAAGCTCTCCAGGAGAGACAAAGC---CAACAGCAACGAGAT 357
DB 124 SerGlnArgAlaProArgPheGluGlyGluAspGlnSerGlnGlnGlnGlnAsp 143
QY 358 AGTCACAGAGGTGCACCGTTTCGATGAGGTGATCTCATTCAGTCCACGCGGTGT 417
DB 144 SerHisGlnIysValArgArgPheAspGluGlyAspLeuIleAlaValProThrGlyVal 163
QY 418 GCTTTCTGGCTCACACACACGACCTGATGTGTGTGCTGCTCTCTCTACTGACACC 477
DB 164 AlaLeuTrpMetTyrAsnAspHisAspThrAspValAlaValSerLeuThrAspThr 183
QY 478 AACAAACACACACAGCTTGATCAGTTCCTCCAGAGGATTCATTTGGCTGGGACACG 537
DB 184 AsnAsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnHis 203
QY 538 GAGCAAGAGTCTTAAGGTACCAACAAAGCAGACAAAGCAGACGAGAGGAGCTTACCA 597
DB 204 GluGlnGluPheLeuArgTyrGlnGln-----GlnSerArgArgSerLeuPro 220
QY 598 TATAGCCCATACAGCCCGCAAGTTCAGCTAGACAAAGAGCGTGAATTTAGCCCTCGA 657
DB 221 TyrSerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArg 240
QY 658 GGCACGACAGCCGACAGAACGACGACGACGACGACGACGACGACGACGACGACGAC 717
DB 241 GlyGlnHisSerArgArgGluArgAlaGlyGlnGlnGlnGlnGlnGlnGlnGlnGln 260
QY 718 TTCAGCGGCTTCACGCGGAGTTTCCTGGAAACAAGCCTTCAGGTTGACGACACAGATA 777
DB 261 PheSerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAspAspArgGlnIle 280
QY 778 GTCAAAACCTAAGAGCGAGACCGGAGGTGAAGAGAGGAGGAGGAGGAGGAGGAGGAG 837
DB 281 LeuGlnAsnLeuArgGlyGluAsnGluSerAspGluGlnGlyAlaIleValThrValArg 300
QY 838 GGAGGCTCAGATCTTGACCCAGACGATAGAAAAGAG-----CGTGGCGAC 882
DB 301 GlyGlyLeuArgIleLeuSerProAspArgLysArgGlnGlnTyrGluArgProAsp 320
QY 883 GAAGAGAGGAATACGATGAAGTGAATATGAATACGATGAAG-----GAT 930

DB 321 GluGluGluGluTyrAspGluAspGluTyrGluTyrAspGluGluGluArgGlnGlnAsp 340
QY 931 AGAAGCGCTGGCAGGGAGACAGACGAGGGGAATGTATTGAAGAGACGATCTGCACC 990
DB 341 ArgArgArgGlyArgGlySerArgGlyArgGlyAsnGlyIleGluGluThrIleCysThr 360
QY 991 GCAAGTCTGCTAAAAGAACATTGGTAGAACACAGATCCCTGACATCTACAAACCTCAAGCT 1050
DB 361 AlaSerValLysLysAsnIleGlyArgAsnArgSerProAspIleTyrAsnProGlnAla 380
QY 1051 GGTTCACCTCAAACTGCCAACGATCTCAACCTTCTAATACTTAGTGGCTTGCAGCTACT 1110
DB 381 GlySerLeuLysThrAlaAsnAspLeuAsnLeuLeuIleLeuArgTrpLeuGlyLeuSer 400
QY 1111 GCTGAATATGGAATCTCTACAGGAATGCAATGTTTGTTCGCTCCTACTACAACCAACGCA 1170
DB 401 AlaGluTyrGlyAsnLeuTyrArgAsnAlaLeuPheValProHisTyrAsnThrAsnAla 420
QY 1171 CACAGCATCATATATCGATTGAGGGACGGGCTCAGTGCAGAGTCTGTCGACAGCACGCG 1230
DB 421 HisSerIleIleTyrAlaLeuArgGlyArgAlaHisValGlnValValAspSerAsnGly 440
QY 1231 AACAGAGTGTACGACGAGGAGCTTCAAGAGGGTCACTGCTTGTGTGTCGACAGAACTTC 1290
DB 441 AsnArgValTyrAspGluGluLeuGlnGluGlyHisValLeuValValProGlnAsnPhe 460
QY 1291 GCGTCGCTGGAAAGTCCAGAGCGAGAACTTCGAAATACGTGGCATTTCAAGACAGACTCA 1350
DB 461 AlaValAlaGlyLysSerGlnSerAspAsnPheGluTyrValAlaPheLysThrAspSer 480
QY 1351 AGCCCCAGCATAGCCAACTCCCGGTGAAACTCCGTATAGATAAACCCTCCGCGAGGAG 1410
DB 481 ArgProSerIleAlaAsnLeuAlaGlyGluAsnSerIleAspAsnLeuProGluGlu 500
QY 1411 GTGGTTGCAATTCATATGCTCCAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1470
DB 501 ValValAlaAsnSerTyrGlyLeuProArgGluGlnAlaArgGlnLeuLysAsnAsn 520
QY 1471 CCCTCAAGTCTTCGTTCCACCGTCTCAGCAGTCTCCAGGGGTGTGGCT 1521
DB 521 ProPheLysPhePheValProProSerGlnGlnSerLeuGlyAlaValAla 537

RESULT 6

QBLKN1_ARAHY PRELIMINARY; PRT; 538 AA.
AC QBLKN1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Allergen Arah3/Arah4.
OS Arachis hypogaea (peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Viquez O.M., Konan K.N., Dodo H.W.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF510854; AAM46958.1; -; Genomic_DNA.
DR HSSP; P04776; IUCX.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin_1; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S_SEED_STORAGE; UNKNOWN 1.
SQ SEQUENCE 538 AA; 61737 MW; 7AABD0D59429709E CRC64;

Alignment Scores:

Pred. No.:	3,38e-162	Length:	538
Score:	2444.00	Matches:	471
Percent Similarity:	93.6%	Conservative:	14
Best Local Similarity:	90.9%	Mismatches:	19
Query Match:	87.9%	Indels:	14
DB:	2	Gaps:	4

US-10-728-323-3 (1-1524) x Q8LKN1 ARAHY (1-538)

Qy	1	CGCGACGAACCGGAGAGAA	CGCGTGCAGTTCCAGCGCCTCAATGCGGACGACCTGAC	60
Db	24	ArgGlnGlnProGluGluAsnAla	CysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	43
Qy	61	AATCGCATTAATCATAGAGGGGGTTTACATT	CAGACTTGGAAACCCCAACACACGAGGATTC	120
Db	44	AsnArgileGluSerGluGlyGlyTyrileGlu	ThrTrpAsnProAsnAsnGlnGluPhe	63
Qy	121	GAATGCGCGCGGTCCCTCTCTCGTTAGTCT	CCGCCCAACGCCCTTCGTAGGCCT	180
Db	64	GluCysAlaGlyValAlaLeuSerArgLeu	VallLeuArgArgAsnAlaLeuArgArgPro	83
Qy	181	TTCTACTCAATGCTCCCGAGAGATCTT	CATCAGCAAGGAGGGGATCTTTGGGTG	240
Db	84	PheTyrSerAsnAlaProGlnGluIlePhe	ileGlnGlyArgGlyTyrPheGlyLeu	103
Qy	241	ATATTCCCTGGTTGCTTAGACACTAT	CAAGAGCCTCACACAAGTGCGTGATCTCAG	300
Db	104	IlePheProGlyCysProSerThrTyrGlu	GluProAlaGlnGlnGlyArgArgHisGln	123
Qy	301	TCCCAAGACCAACCAAGACGCTCT	CAAGGAGAAGCAACCAAGCAACAGACAGATGT	360
Db	124	SerGlnArgProProArgArgPheGln	GlyAspGlnSerGlnGlnGlnAspSer	143
Qy	361	CACCAAGGTCGACCGTTTGCATGAGGGT	GATCTCATTTGAGTTCACCGGTGTGCT	420
Db	144	HisGlnLysValIleArgPheAspGlu	GlyAspLeuIleAlaValProThrGlyValAla	163
Qy	421	TTCTGCTCTCAACACGACACGACACT	GATGTGTGCTGTTCTTCTTACTGACACCAAC	480
Db	164	PheTrpMetCysAsnAspHisAspThr	AspValValAlaValSerLeuThrAspThrAsn	183
Qy	481	AACAACGCAACCAACGCTTGATCAGT	TCCCCAGGAGATTCAATTTGGCTGGGAACACG	540
Db	184	AsnAsnAspAsnGlnLeuAspGlnPhe	ProArgArgPheAsnLeuAlaGlyAsnHisGlu	203
Qy	541	CAAGAGTTCTTAAGTACCAAGCAACAA	AGCAGACAAAGCAGACGAAGAAGCTTACCATAT	600
Db	204	GlnGluPheLeuArgTyrGlnGln	-----GlnSerArgArgSerLeuProTyr	220
Qy	601	AGCCCATACAGCCGCAAGTCCAGCT	TAGACAAGAGCGTGAATTTAGCCCTCCAGCA	660
Db	221	SerProTyrSerProGlnThrGlnPro	lysglnGluAspArgGluPheSerProArgGly	240
Qy	661	CAGCACAGCCGAGAAACGACGAGCA	GAAGAAGAAACCAAGGTGGAAACATCTTTC	720
Db	241	GlnHisGlyArgArgGluArgAlaGly	GlnGlnGluGlnGluAsnGluGlyGlyAsnIlePhe	260
Qy	721	AGCGGCTTCACGCGGAGTTCTG	TGAAACAACGCTTCAGGTTGACGACGACAGATAGTG	780
Db	261	SerGlyPheThrProGluPheLeuAla	GlnAlaPheGlnValAspAspArgGlnIleLeu	280
Qy	781	CAAAACCTAAGAGCGGACGACGAG	TCAAGAGGAGCCATTGTGACAGTGGGGA	840
Db	281	GlnAsnLeuArgGlyGlyLeuAsnGlu	SerAspGluGlnGlyAlaIleValThrValArgGly	300
Qy	841	GGCCTCAGAAATCTTGAGCCCGCAT	AGAAAGAGA-----CGTCCCGCAGAA	885
Db	301	GlyLeuArgIleLeuSerProAspArg	LysArgArgGlnGlnTyrGluArgProAspGlu	320
Qy	886	GAAGAGGAATACGATGAAGATGAT	TGATGATGATGAAGAG-----GATAGA	933
Db	321	GluGluGlyTyrAspGluAspGlyTyr	GlnTyrAspGluGluArgGlnGlnAspArg	340

Qy	934	AGCGTGGCAGGGGACGACAGCAGGGGGAGTGTATGAAGACAGAGTCTCCACCGCA	933
Db	341	ArgArgGlyArgGlySerArgGlySerGlyAsnGlyLeuGluGluThrIleCysThrAla	360
Qy	994	AGTGCTAAAAGAACATTGGTAGAAACAGATCCCTGCACATCTACAAACCTCAAGGTGGT	1053
Db	361	SerPheIysIysAsnIleGlyAArgAsnArgSerProAspIleTyrAsnProGlnAlaGly	380
Qy	1054	TCACCTCAAAACTGCCAACGAT-----CTCAACTCTTAATATCTTAGTGGCTTGGACCT	1107
Db	381	SerLeuIysThrAlaAsnGluLeuLeuAsnLeuLeuIleLeuArgTyrLeuGlyLeu	400
Qy	1108	AGTGCTGAATATGGAATCTCTACAGGAATGCATTGTTGTCCTCACTACACCAACAAC	1167
Db	401	SerAlaGluTyrGlyAsnLeuTyrArgAsnAlaLeuPheValProHisTyrAsnThrAsn	420
Qy	1168	GCACACAGCATCATATCGATTGAGGGGACGGGCTCACGTGCAAGTCGTGGACAGCAAC	1227
Db	421	AlaHisSerIleIleTyrAlaLeuArgGlyArgAlaHisValGlnValValAspSerAsn	440
Qy	1228	GGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTCACGTGTTGTCGCCACAGAAC	1287
Db	441	GlyAspArgValPheAspGluLeuLeuGlnGluGlyHisValLeuValValProGlnAsn	460
Qy	1288	TTGCGCTGCTCGAAAGTCCCAAGAGGGAGAACTTCGAATACGTGGCATTCAGACAGAC	1347
Db	461	PheAlaValAlaGlyIysSerGlnSerGluAsnPheGluTyrValAlaPheIysThrAsp	480
Qy	1348	TCAGGCCCGACATAGCCCAACTCGCGGTGAAACTCCGTCATAGATACCTGCCGGAG	1407
Db	481	SerArgProSerIleAlaAsnLeuAlaGlyGluAsnSerPheIleAspAsnLeuProGlu	500
Qy	1408	GAGTGTGGTGCAAATTATATGGCTCTCAAGGGGACGAGGAGGACGAGCTTAAGAACAAAC	1467
Db	501	GluValValAlaAsnSerTyrGlyLeuProArgGluGlnAlaArgGlnLeuIysAsnAsn	520
Qy	1468	AAOCCCTTCAAGTTCCTCGTTCACCGCTCTCAGACAGTCTCCGAGGGGTGTGGCT	1521
Db	521	AsnProPheIysPhePheValProProSerGlyGlnSerLeuArgAlaValAla	538

RESULT 7

Q6T2T4_ARAHY
 ID Q6T2T4_ARAHY PRELIMINARY; PRT; 536 AA.
 AC Q6T2T4;
 DT 05-JUL-2004 (T-EMBLrel. 27, Created)
 DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
 DE Storage protein.
 OS Arachis hypogaea (peanut).
 OC Arachis hypogaea;
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae
 OC Arachis.
 OX NCBI TaxID=3818;

```

Alignment Scores:
Pred. No.:      1,22e-161      Length:      536
Score:          2436.00        Matches:     468
Percent Similarity: 93.6%      Conservative: 15
Best Local Similarity: 90.7%   Mismatches:  21

```

Query Match:	87.7%	Indels:	12
DB:	2	Gaps:	3
US-10-728-323-3 (1-1524) x Q6T2T4_ARAHY (1-536)			
QY 1	CGGCAGCAACCGGAGGAGACGGTCCAGTTCACCGCTCAATCCGCGAGACCTGAC	60	
DB 24	ArgGlnGlnProGluGlnAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	43	
QY 61	AATCGCATTAATPCAGAGCGCGTTACATTGAGACTTGGAACCCCAACACGAGGATTC	120	
DB 44	AsnArgIleGluSerGluGlyGlyTyriIeGluThrTriAsnProAsnAsnGlnGluPhe	63	
QY 121	GAATGCGCGCGTCCCTCTCTCGTTAGTCTCTCGCGGCACACCCCTTCGTAGGCTT	180	
DB 64	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro	83	
QY 181	TTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGATCTTGGTTG	240	
DB 84	PheTyriSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgAlaTyriPheGlyLeu	103	
QY 241	ATATTCCTCGTGTCTCTAGACACTATGAAGACCTCACACAAGGTCTGTCGATCTCAG	300	
DB 104	IlePheLeuGlyCysProSerThrTyriGluGluProAlaGlnGlnGlyArgArgHisGln	123	
QY 301	TCCCAAGACCAACAGACGCTTCCAGGAGAAAGCAACCAAGCAACGAGATAGT	360	
DB 124	SerGlnArgProProArgPheGlnGlyGlnAspGlnSerGlnGlnGlnGlnAspSer	143	
QY 361	CACAGAGGTGCACCGTTTCGATGAGGTGATCTCATTCAGTTCCTCCACCGGTGTGCT	420	
DB 144	HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla	163	
QY 421	TTCTGCTCTPACACGACCAACGACACTGATGTGTGCTGCTTCTCTTACTGACCAAC	480	
DB 164	PheTriMetTyriSerAsnAspHisAspThrAspValAlaValSerLeuThrAspThrAsn	183	
QY 481	AACAACGACACACGCTGATTCAGTTCCTCCAGGAGATTCATTTGGCTGGACACGGAG	540	
DB 184	AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHisGlu	203	
QY 541	CAAGAGTCTTAAAGTACACGACCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT	600	
DB 204	GlnGluPheLeuArgTyriGlnGln-----GlnSerArgArgSerLeuProTyri	220	
QY 601	AGCCCATACGCGCCGAAAGTACGCTAGACAAAGAGAGCGTGAATTTAGCCCTCGAGGA	660	
DB 221	SerProTyriSerProGlnThrGlnProLysGlnGluAspArgGluPheSerProArgGly	240	
QY 661	CAGCAGCGCCGACAGACGAGCAGGACAGACAGAGAGAGAGAGAGAGAGATCTTC	720	
DB 241	GlnHisGlyArgArgGluArgAlaGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnPhe	260	
QY 721	AGCGGCTTACGCGCGAGTTCCTGGAAACAGCTTCAGGTTCCAGCAGACAGACAGATAGT	780	
DB 261	SerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAspAspArgGlnIleLeu	280	
QY 781	CAAAACCTAAGAGCGGACCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	840	
DB 281	GlnAsnLeuArgGlyGluAsnGluSerAspGluGlnGlyAlaIleValThrValArgGly	300	
QY 841	GGCTCAGATCTTGAGCCAGACAGATAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG	885	
DB 301	GlyLeuArgIleLeuSerProAspArgLysArgGlnGlnTyriGluArgProAspGlu	320	
QY 886	GAAGAGGAATACGATGAAGATGAATATGAATACGATGAAGAG-----GATAGA	933	
DB 321	GluGluGluTyriAspGluAspGluTyriGluTyriAspGluGluGluGlnGlnAspArg	340	
QY 934	AGCGGTGGCAGGGAAGCAGAGGCGAGGGAATGGTATTGAAGAGACGATCTCGACCGCA	993	
DB 341	ArgArgGlyArgGlySerArgGlySerGlyAsnGlyIleGluGluThrIleCysThrAla	360	

QY 994	AGTGCTGTAAGAAACATTGGTAGAAACAGATCCCTCGATCATCAACCCCTCAAGCTGCT	1053	
DB 361	SerPheLysLysAsnIleGlyArgAsnArgSerProAspIleTyriAsnProGlnAlaGly	380	
QY 1054	TCACCTCAAAACCTGCCAACGACTCAACCTTCTTAATACTTAGTGGCTTTGGACCTAGTGT	1113	
DB 381	SerLeuLysThrAlaAsnGluLeuAsnLeuLeuIleLeuArgTriPheGlyLeuSerAla	400	
QY 1114	GAATATGGAATCTCTACAGGAATCGATTGTTTGTTCGCTCACTACACCAACGACACAC	1173	
DB 401	GluTyriGlyAsnLeuTyriArgAsnAlaLeuPheValProHisTyriAsnThrAsnAlaHis	420	
QY 1174	ACATCATATATCCGATTGAGGGACGGCTCAGTCGCAAGTCGTCGACAGCAACGCAAC	1233	
DB 421	SerIleIleTyriAlaLeuArgGlyArgAlaHisValGlnValValAspSerAsnGlyAsp	440	
QY 1234	AGAGTGTACGACGAGGAGCTTCAAGAGGCTCACGTCCTTGTGTGGTCCACAGAACTTCGCC	1293	
DB 441	ArgValPheAspGluGluLeuGlnGlyHisValLeuValValProGlnAsnPheAla	460	
QY 1294	GTCGCTGGAATCCCGACGAGCAGAACTTCGAATACGTGGCATTCAAGACAGACTCAAGG	1353	
DB 461	ValAlaGlyLysSerGlnSerGluAsnPheGluTyriValAlaPheLysThrAspSerArg	480	
QY 1354	CCAGCATAGCAACCTCGCGGTGAAACTCCGTCATAGATAACCTGCCGAGGAGGTG	1413	
DB 481	ProSerIleAlaAsnLeuAlaGlyGluAsnSerPheIleAspAsnLeuProGluVal	500	
QY 1414	GTTGCAAAATTCATATGGCTCCAAAGGGAGCAGCAAGCAGCTTAAGAACCAACCC	1473	
DB 501	ValAlaAsnSerTyriGlyLeuProArgGluGlnAlaArgGlnLeuLysAsnAsnPro	520	
QY 1474	TTCAAGTCTTCGTCACCGCTCTCAGCAGTCTCCGAGGCTGTGGCT	1521	
DB 521	PheLysPhePheValProProSerGluGlnSerLeuArgAlaValAla	536	
RESULT 8			
Q9SQH7_ARAHY	PRELIMINARY;	PRT;	530 AA.
ID	Q9SQH7;		
AC	Q9SQH7;		
DT	01-MAY-2000 (Tremblrel. 13, Created)		
DT	01-MAY-2000 (Tremblrel. 13, Last sequence update)		
DT	01-JUN-2003 (Tremblrel. 24, Last annotation update)		
DE	Glycinin.		
GN	Name=Arach4;		
OS	Arachis hypogaea (Peanut).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;		
OC	rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;		
OC	Arachis.		
OX	NCBI_TaxID=3818;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=Virginia;		
RX	MEDLINE=99406463; PubMed=10474031; DOI=10.1159/000024203;		
RA	Kleber-Janke T.; Cramer R.; Appenzeller U.; Schlaak M.; Becker W. -M.;		
RT	"Selective cloning of peanut allergens, including profilin and 2S		
RL	albumins, by phage display technology.";		
RL	Int. Arch. Allergy Immunol. 119:265-274 (1999).		
DR	EMBL; AF086821; AAD47382.1; -; mRNA.		
DR	HSSP; P04776; IUCX.		
DR	GO; GO:0045735; F.nutrient reservoir activity; IEA.		
DR	InterPro; IPR006045; Cupin.		
DR	InterPro; IPR007113; Cupin_region.		
DR	InterPro; IPR006044; Seedstore_11s.		
DR	Pfam; PF00190; Cupin_1; 2.		
DR	PRINTS; PR00439; 11SGLOBULIN.		
SQ	SEQUENCE 530 AA; 61011 MW; 864E349A31642A32 CRC64;		
Alignment Scores:			
Pred. No.:	1.36e-160	Length:	530
Score:	2421.00	Matches:	469
Percent Similarity:	94.1%	Conservative:	9

Best Local Similarity: 92.3%
Query Match: 87.1%
DB: Indels: 2
Gaps: 0

US-10-728-323-3 (1-1524) x Q9SQH7_ARAHY (1-530)

```

QY 1 CGGAGCAACCGGAGGAGAACGGTGCAGTTCACGCGCTCAATGCGGAGAGACTGAC 60
DB 24 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
QY 61 AATCCGATTGAATCAGAGCGCGGTACATTGAGACTTGGAACCCCAACACAGGAGTTC 120
DB 44 AsnArgileGluSerGluGlyTyriLeGluThrTrpAsnProAsnAsnGlnGluPhe 63
QY 121 GAATGCCCGCGGTGCCCTCTCTCGCTTAGTTCCTCCGCGCAACGCGCTTCGTAGGCT 180
DB 64 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro 83
QY 181 TTCTACTCCAATGCTCCCCAGGAGATCTTCATCCAGCAAGGAGGATACCTTCGGTTG 240
DB 84 PheTyrSerAsnAlaProGlnGluPheIleGlnGlnGlyArgGlyTyrPheGlyLeu 103
QY 241 ATATTCCCTGCTGCTCCTAGACACTATGAGAGCCTCACAACAAGTCTGATCTCAG 300
DB 104 IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgArgTyrGln 123
QY 301 TCCCAAGACCAACGAGCTCTCAAGGAGAGAACCAAGCAACAGCAACGAGATAGT 360
DB 124 SerGlnArgProProArgArgLeuGlnGluAspGlnSerGlnGlnGlnAspSer 143
QY 361 CACAGAAAGTGACCGTTTCATGAGGTGATCTCATTCAGTTCCTCCACCGGTGCT 420
DB 144 HisGlnLysValHisArgPheAsnGluGlyAspLeuIleAlaValProThrGlyValAla 163
QY 421 TTCTGCTCTACACGACCAACACTGATGTGTGCTGTTCTTCTTACTACACCAAC 480
DB 164 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 183
QY 481 AACACGACCAACGAGTTCAGTTCCTCCAGGAGATTCATTTGCTGGGAAACGAG 540
DB 184 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHisGlu 203
QY 541 CAAGAGTTCTTAAGTTACAGCAACAAAGACAGCAAGACGAGAGCTTACCATAT 600
DB 204 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 223
QY 601 AGCCCATACAGCCGCAAGTACGCTAGACAGAGCGGTGAATTTAGCCCTCAGGA 660
DB 224 SerProTyrSerProHisSerArgProArgArgGluGluArgGluPheArgProArgGly 243
QY 661 CAGCAGCGCGCAGAGACGAGCAAGCAAGAAAGAAAGCAAGGTGGAACATCTTC 720
DB 244 GlnHisSerArgArgGluArgAlaGlyGlnGlnGluGluAspGluGlyGlyAsnIlePhe 263
QY 721 AGCGGCTTCCCGGAGTTCCTGGAACAAAGCTTCAGTTGACGACAGACAGATG 780
DB 264 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 283
QY 781 CAAAACCTAAGCGCGCAGACCGAGTGAAGAGAGGAGCGCATTTGACAGTGGGGA 840
DB 284 GlnAsnLeuTrpGlyGluAsnGluSerGluGluGluGlyAlaIleValThrValArgGly 303
QY 841 GCGCTCAGAAATCTGAGCCAGATAGAAAGACGCTGCCGACGAAAGAGAGAAATACCAT 900
DB 304 GlyLeuArgIleLeuSerProAspGlyThrArgGlyAlaAspGluGluGluGluThrAsp 323
QY 901 GAAGATGAATATGATACGATCAAGAGGATAGAACGGTGGCAGGGGAGCAGAGCAGG 960
DB 324 GluAspGlnTyrGluTyrHisGluGlnAspGlyArgGlyArgGlySerArgGlyGly 343
QY 961 GGGATGGTATTGAAGACGAGTTCGACCCGCAAGTGTAAAGAACATCTGTTGAAC 1020
DB 344 GlyAsnGlyIleGluGluThrIleCysThrAlaCysValLysLysAsnIleGlyGlyAsn 363

```

```

QY 1021 AGATCCCTGACATCTACAACCCCTCAAGC-TGGTTTCACTCAAAAAGTGCACGATCTCAA 1079
DB 364 ArgSerProHisIleTyrAspProGlnArgTrpPheThrGlnAsnCys-HisAspLeuAs 383
QY 1080 CTTCTTAATACTAGGTGGCTTGGACCTAGTGTCTGAATATGGAATATCTCTACAGGAATGC 1139
DB 383 nLeuLeuIleLeuArgTrpLeuGlyLeuSerAlaGluTyrGlyAsnLeuTyrArgAsnAl 403
QY 1140 ATTGTTTTCGCTCACTCAACACCAACGACACAGCATCATATATCGATTTCAGGGGACG 1199
DB 403 aleuPheValProHisTyrAsnThrAsnAlaHisSerIleIleTyrAlaLeuArgGlyAr 423
QY 1200 GGCTCACGTCGCAAGTCTGCGACGACCAACGCAACAGAGTGTACGACGAGGACTTCAACA 1259
DB 423 gAlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnG 443
QY 1260 GGGTCACGTCGTTGTGTGGCCACAGAACTTCGCCCTCGCTGGAAAGTCTCCAGAGGAGAA 1319
DB 443 uGlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAs 463
QY 1320 CTTGGAATACGTCGTCATTCAGACAGACTCAAGGCCACAGCAGTAGCAACCTCGCGGTCA 1379
DB 463 nPheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnPheAlaGlyG 483
QY 1380 AAATCCCTCATAGATAACCTCGCGGAGAGGTGTTGCAAAATTCATATGGCCTCCAAAG 1439
DB 483 uAsnSerPheIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuProAr 503
QY 1440 GGAGCAGCAGCAGCAGCTTAAGAACAAACACCCCTTCAAGTTCCTGTTCCACCGTCTCA 1499
DB 503 gGluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProProPheG 523
QY 1500 GCAGTCTCCGAGGCGTGTGGCT 1521
DB 523 mGlnSerProArgAlaValAla 530

```

RESULT 9

```

Q61WGS_ARAHY PRELIMINARY; PRT; 510 AA.
ID Q61WGS5;
AC Q61WGS5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Glycinin (Peanut).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RA Kang I.-H., Gallo-Meagher M.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY618460; AAT39430.1; -, mRNA.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin 1; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S_SEED_STORAGE; UNKNOWN_1.
FT NON TER 1
SQ SEQUENCE 510 AA; 58061 MW; F6569EF13D1A12BD CRC64;

```

Alignment Scores:

```

Pred. No.: 1.45e-119 Length: 510
Score: 1834.00 Matches: 364
Percent Similarity: 78.6% Conservative: 43
Best Local Similarity: 70.3% Mismatches: 71
Query Match: 66.0% Indels: 40
DB: Gaps: 2

```

US-10-728-323-3 (1-1524) x Q61W5_ARAHY (1-510)

QY 1 CGGACGAAACCGGAGGAGAACCGGTGCCAGTTCACGGCTCAATCGGAGAGACCTGAC 60
 DB |||||
 DB 22 ArgGlnGlyGluGluAsnGluCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 41
 QY 61 AATCGATTGAATCAGAGCGCGGTACATTGAGACTTGGAAACCCCAACACGAGGAGTTC 120
 DB |||||
 DB 42 AsnArgIleGluSerGluGlyGlyTyrlleGluThrTrpAsnProAsnGlnGluPhe 61
 QY 121 GAATCGCGCGGCTGCCCTCTCCGTTAGTCTCCGCGCAACCCCTTCGTAGCCCT 180
 DB |||||
 DB 62 GlnCysAlaGlyValAlaLeuSerArgThrValLeuArgAsnAlaLeuArgPro 81
 QY 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGATACTTTGGTTG 240
 DB |||||
 DB 82 PheTy-SerAsnAlaProLeuGluIleTyrlleValGlnGlnGlySerGlyTyrllePheGlyLeu 101
 QY 241 ATATTCCCTGGTTGCTTAGACACTATGAGAGCCTTCACACAAAGTTCGTGCTCAG 300
 DB |||||
 DB 102 IlePheProGlyCysProSerThrTyrlleGluGluProAlaGlnGluGlyArgTyrlleGln 121
 QY 301 TCCCAAGACCAACAGAGCTCTCCAA-----GGAGAAAGACCAAGCCAAACAGCAACGA 354
 DB |||||
 DB 122 SerGlnLysProSerArgPheGlnValGlyGlnAspProSerGlnGlnGln 141
 QY 355 GATAGTCACAGAGGTGCACCTTTCGATGAGGGTGATCTCATTCAGTTCCTCCACCGGT 414
 DB |||||
 DB 142 AspSerHisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGly 161
 QY 415 GTTGCTTTCGGCTCTACACGACCCAGCACATGATGTTGTTCTCTCTTCTACTGAC 474
 DB |||||
 DB 162 ValAlaPheTrpMetTyrlleAsnAspGluAspThrAspValThrValThrLeuSerAsp 181
 QY 475 ACCAAACAAACACACACACAGCTTGATCAGTCTCCCGAGGAGATTCAATTTGGCTGGGAAC 534
 DB |||||
 DB 182 ThrSerSerIleHisGlnGlnLeuAspGlnPheProArgPheTyrlleAlaGlyAsn 201
 QY 535 ACGGACGAGAGTTCCTTAAGGTACGAGCAACAAAGCAGACAAAGCAGACGAGAAAGCTTA 594
 DB |||||
 DB 202 GlnGluGlnGluPheLeuArgTyrlleGlnGlnGlnGlySer----- 215
 QY 595 CCATATAGCCCATACAGCCCGCAAGTCAAGCTCAGACGAGGCTGGAATTTAGCCCT 654
 DB |||||
 DB 216 -----ArgProHisTyrlleArgGlnIleSerPro 224
 QY 655 CGAGGACAGCAGCCCGCAGAGAACGAGCAGAGCAAGAGAAAGAAAGAGGTGGAAC 714
 DB |||||
 DB 225 ArgValArg-----GlyAspGluGlnGluAsnGluGlySerAsn 237
 QY 715 ATCTTCAGCGGCTTCAGCCCGAGTTCCTGGAAACAAGCCTTCAGGTTGACGACAGACAG 774
 DB |||||
 DB 238 IlePheSerGlyPheAlaGlnGluPheLeuGlnHisAlaPheGlnVal---AspArgGln 256
 QY 775 ATAGTCAAACCTTAAGAGGCGAGACCGAGAGTGAAGAGAGGAGGACATTGTGACAGTG 834
 DB |||||
 DB 257 ThrValGlnAsnLeuArgGlyGluAsnGluArgGluGluGlnGlyAlaIleValThrVal 276
 QY 835 AGGGGAGGCTCAGAAATCTTCGAGCCCAAGATAGAAAGAGAGCTGCC----- 879
 DB |||||
 DB 277 LysGlyGlyLeuArgIleLeuSerProAspGluGluAspGluSerSerArgSerProPro 296
 QY 880 GACGAAGAAGGATACGATGAGATGAATGAATATGAATACGATGAGAGGATAGAGCGCT 939
 DB |||||
 DB 297 SerArgArgGluGluPheAspGluAspArg-----SerArgProGlnGlnArg 312
 QY 940 GGACAG-----GGAAGCAGGCGAGGGGAAATGGTATTCGAAGAGAGCTTCGACC 990
 DB |||||
 DB 313 GlyLysTyrlleAspGluAsnArgArgGlyTyrlleAsnGlyIleGluGluThrIleCysSer 332
 QY 991 GCAAGTCTAAAAGAAATTTGGTAGAAACAGATCCCTGACATCTACACCTCAAGCT 1050
 DB |||||
 DB 333 AlaSerValLysLysAsnLeuLeuArgSerSerAsnProAspIleTyrlleAsnProGlnAla 352

QY 1051 GGTTCACCTCAAACTGCCAACGATCTCAACCTTCTAATACTTAGGTGGCTTGACCTACT 1110
 DB 353 GlySerLeuArgSerValAsnGlnLeuAspLeuProIleLeuGlyTyrlleLeuGlyLeuSer 372
 QY 1111 GGTGAATATGGAATCTCTACAGGAATGATTGTGTTCGCTCACTACAACCAACGCA 1170
 DB |||||
 DB 373 AlaGlnHisGlyThrIleTyrlleArgAsnAlaMetPheValProHisTyrlleLeuAsnAla 392
 QY 1171 CACGACATCATATATCGATTGAGGGAGCGGTCTACGTCGAAGTCGTGGACAGCAACGCGC 1230
 DB |||||
 DB 393 HisThrIleValValAlaLeuAsnGlyArgAlaHisValGlnValValAspSerAsnGly 412
 QY 1231 AACAGAGTGTACGACGAGAGCTTCAAGAGGCTCAGCTTCTGTGTGGTCCACAGAACTTC 1290
 DB |||||
 DB 413 AsnArgValTyrlleAspGluGluLeuGlnGlyHisValLeuValValProGlnAsnPhe 432
 QY 1291 GCGCTCGCTGGAAGTCCCGAGAGCAGAACTTCGAATACGTGGCATTTCAAGACAGACTCA 1350
 DB |||||
 DB 433 AlaValAlaAlaLysAlaGlnSerGluAsnTyrlleGluTyrlleLeuAlaPheLysThrAspSer 452
 QY 1351 AGGCCAGCATAGCCAACTCCCGGTGAAACTCCGTCTATAGATAAACCCTCGCGGAGGAG 1410
 DB |||||
 DB 453 ArgProSerIleAlaAsnLeuAlaGlyGluAsnSerIleIleAspAsnLeuProGluGlu 472
 QY 1411 GTGGTTCGAATTCATATGCCCTCCAAAGGAGGAGCAGCAAGCAGCTTAAGAACACAAAC 1470
 DB |||||
 DB 473 ValValAlaAsnSerTyrlleArgLeuProArgGluGlnAlaArgGlnLeuLysAsnAsn 492
 QY 1471 CCCTTCAAGTTCCTGCTCCACCG---TCTCAGCAGTCTCCGAGGCTGTGGCT 1521
 DB |||||
 DB 493 ProPheLysPhePheValProPheAspHisGlnSerMetArgGluValAla 510

RESULT 10

Q647H2 ARAHY
 ID Q647H2 ARAHY PRELIMINARY; PRT; 484 AA.
 AC Q647H2;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Arachis hly-3.
 OS Arachis hypogaea (Peanut).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC Rosids; eurosids 1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
 OC Arachis.
 OX NCBI_TaxID=3818;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Van Y.S., Lin X.D., Zhang Y.S., Wang L., Wu K., Huang S.Z.;
 RT "Isolation of peanut genes encoding arachins and conglutins by
 RT expressed sequence tags.";
 RL Plant Sci. 169:439-445 (2005).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Van Y., Wang L., Huang S.;
 RT "cDNA clone of peanut seed storage protein gene.";
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY722687; AAU21492.1; -; mRNA.
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR006045; Cupin.
 DR InterPro; IPR007113; Cupin region.
 DR InterPro; IPR006044; Seedstore_11s.
 DR Pfam; PF00190; Cupin_1; 2.
 DR PRINTS; PR00439; 11SGLOBULIN.
 DR PROSITE; PS00305; 11S SEED STORAGE; UNKNOWN.1.
 SQ SEQUENCE 484 AA; 54569 MW; 5A3E950752E89D2D CRC64;

Alignment Scores: 8.9e-107 Length: 484
 Pred. No.: 1651.00 Matches: 337
 Score: 1651.00
 Percent Similarity: 75.2% Conservative: 46
 Best Local Similarity: 66.2% Mismatches: 76

Query Match: 59.4% Indels: 50
DB: 2 Gaps: 8
US-10-728-323-3 (1-1524) x Q647H2_ARAHY (1-484)

QY 1 CGGCAGCAACCGGAGAGAACCGGTGCTCCAGGCGCTCAATGCGGAGAGACTGAC 60
DB 24 ArgGlnGlnGlyGluGluAsnGluCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
QY 61 AATCGCATTAATACAGAGCGCGGTACATTGAGACTTGGAACCCCAACACAGAGAGTTC 120
DB 44 AsnCysIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 63
QY 121 GAATGCGCGCGGTGCTCTCTCGCTTAGTCTCTCCGCGCAACGCGCTTCTGAGGCT 180
DB 64 GlnCysAlaGlyValAlaLeuSerArgPheValLeuArgArgAsnAlaLeuArgPro 83
QY 181 TTCTACTCAATGCTCCCGAGAGATCTTCATCCAGCAAGGAGGGGATACCTTGGGTG 240
DB 84 PheTyrSerAsnAlaProGlnGluIlePheIleTyrGlnGlySerGlyTyrPheGlyLeu 103
QY 241 ATATTCCCTGGTGTCTTAGACACTATGAAGAGCTTCACACAAAGTCTCGATCTCAG 300
DB 104 IlePheProGlyCysProGlyThrPheGluGluPro---IleGlnGly---SerGluGln 121
QY 301 TCCCAAGACCAACGAGCGTCTCCAGGAGAGAGCAAAAGCCACAGCAACGAGATAGT 360
DB 122 PheGlnArgProSerArgHisPheGlnGlyGlnAspGlnSerGlnArgProLeuAspThr 141
QY 361 CACAGAGGTGACCGTTTCAGATGAGGTGATCTCATTTGCGATGCCACCGGTGTGCT 420
DB 142 HisGlnLysValHisGlyPheArgGluGlyAspLeuIleAlaValProHisGlyValAla 161
QY 421 TTCTGGCTCTACAAACACACGACACTGATGTTGTCGTTCTTCTTACTGACACCAAC 480
DB 162 PheTrpIleTyrAsnAspGlnAspThrAspValValAlaIleSerValLeuHisThrAsn 181
QY 481 AACACGACCAACGAGCTTTCAGTTCCCGCAGGAGATTCAATTTGGCTGGCAACGAG 540
DB 182 SerLeuHisAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyLysGlnGlu 201
QY 541 CAGAGTTCTTAAGTTACAGCAACAAAGC---AGACAAGAGCAGCAAGAGAGCTTACCA 597
DB 202 GlnGluPheLeuArgTyrGlnGlnArgSerGlyArgGlnSer----- 215
QY 598 TATAGCCCATACAGCCCGCAAGTCAAGCTAGACAGAGAGAGCGTGCAATTTAGCCCTCGA 657
DB 216 -----ProLysGlyGluGluGlnGlu----- 222
QY 658 GGACAGCACAGCGCCAGAGAACGAGCAGGACAGAGAGAGAGAAACGAAAGTGGAACATC 717
DB 223 -----GlnGluGlnGluAsnGluGlyGlyAsnVal 232
QY 718 TTCAGCGCTTCACGCGGAGGTTCTGGAAACAGCCCTTCAGTTGACGACAGACAGATA 777
DB 233 PheSerGlyPheSerThrGluPheLeuSerHisGlyPheGlnValAsnGlu---AspIle 251
QY 778 GTCCAAACCTAAGAGCGGACCGAGAGTGAAGAGAGGAGGAGCCATCTGACAGCTGAGG 837
DB 252 ValArgAsnLeuArgGlyGluAsnGluArgGluGluGlnGlnGlyValAlaIleValThrValLys 271
QY 838 GGAGGCTTCAGATCTTGAAGCCAGCATAGAAAGAGAGAGCGTCCGACGAAGAGGAATAC 897
DB 272 GlyGlyLeuSerIleLeuValProPro----- 280
QY 898 GATGAAGATGAATATGAATACGATGAAGAGGATAGAAAGCGGTGGCGGGAACGAGAGC 957
DB 281 -----GluTrpArgGlnSerTyrGlnGlnProGlyArgGlyAspLysAsp 295
QY 958 AGGGGAATGGTATTGAAGAGAGATCTCGACCGCAAGTGCTTAAAGAAAGACATGGTAGA 1017
DB 296 PheAsnAsnGlyLeuGluThrIleCysThrAlaThrValLysMetAsnIleGlyLys 315

QY 1018 AACAGATCCCTGTGACATCTACAAACCTCAAGCTGTGTTCTACTCAAAACATGCCAACGATCTC 1077
DB 316 SerThrSerAlaAspIleTyrAsnProGlnAlaGlySerValArgThrValAsnGlnLeu 335
QY 1078 AACCTTCTTAATACCTTAGGTGGCTTGACCTAGTGTGAATATGGAATCTCTACAGGAAT 1137
DB 336 AspLeuProIleLeuAsnArgLeuGlyLeuSerAlaGluTyrGlySerIleHisArgAsp 355
QY 1138 GCATTGTTTGTGCTCACTACAAACCAACGACACAGCATCATATATCGATTGAGGGA 1197
DB 356 AlaMetPheValProHisTyrAsnMetAsnAlaAsnSerMetIleTyrAlaLeuHisGly 375
QY 1198 CGGCTCAGCTCAGTCAAGTGTGGACAGCAACGCAACAGAGTGTACGACGAGAGCTTCAA 1257
DB 376 GlyAlaHisValGlnValValAspCysAsnGlyAsnArgValPheAspGluGluLeuGln 395
QY 1258 GAGGCTCAGCTGCTGTGTCACAGAACTTCGCGCTGCTGGAAAGTCCCAGAGCGAG 1317
DB 396 GluGlyGlnSerLeuValValProGlnAsnPheAlaValAlaAlaLysSerGlnSerGlu 415
QY 1318 AACTTCGAATACGTGGCATTCGAAGACAGACTCAAGGCCGAGCATAGCCACCTCGCGGT 1377
DB 416 HisPheLeuTyrValAlaPheLysThrAsnSerArgAlaSerIleSerAsnLeuAlaGly 435
QY 1378 GAAACTCCGTCAATAGATAACCTCGGAGGAGGTGGTTCGAAATTCATATGCGCTCCAA 1437
DB 436 LysAsnSerTyrMetTrpAsnLeuProGluAspValValAlaAsnSerTyrGlyLeuGln 455
QY 1438 AGGAGAGGAGGAGGAGCGCTTAAGAACACACACCCCTCAAGCTTCCTGCTCCACCG--- 1494
DB 456 TyrGluGlnAlaArgGlnLeuLysAsnAsnProPheThrPheLeuValProProGln 475
QY 1495 TCTCAGCAGCTCTCCGAGGCTGTGCT 1521
DB 476 AspSerGlnMetIleArgThrValAla 484

RESULT 11

GLY3 SOYBN
ID GLY3 SOYBN STANDARD; PRT; 481 AA.
AC P11828;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DE 10-MAY-2005 (Rel. 47, Last annotation update)
DE Glycinin G3 precursor [Contains: Glycinin A subunit; Glycinin B subunit].
GN Name=GX3;
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
OC Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=cv. Dare; TISSUE=Leaf;
RX MEDLINE=89296500; PubMed=2740231;
RA Cho T.-J., Nielsen N.C.;
RT "The glycinin Gy3 gene from soybean.";
RL Nucleic Acids Res. 17:4388-4388(1989).
RN [2]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=92393391; PubMed=2485233; DOI=10.1105/tpc.1.3.313;
RA Nielsen N.C., Dickinson C.D., Cho T.-J., Thanh V.H., Scallion B.J.,
RA Fischer R.L., Sims T.L., Drews G.N., Goldberg R.B.;
RT "Characterization of the glycinin gene family in soybean.";
RL Plant Cell 1:313-328(1989).
CC -!- FUNCTION: Glycinin is the major seed storage protein of soybean.
CC -!- SUBUNIT: Hexamer; each subunit is composed of an acidic and a basic chain derived from a single precursor and linked by a disulfide bond.
CC -!- SIMILARITY: Belongs to the 11S seed storage protein (globulins) family.
CC -----

DE Glycinin G2 precursor [Contains: Glycinin A2 subunit; Glycinin B1a subunit].

DE Name=G2;

GN Glycine max (Soybean).

OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; OC Glycine.

OX NCBI_TaxID=3847;

RN [1]_TaxID=3847;

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=cv. Dare; TISSUE=Leaf;

RX MEDLINE=92393391; PubMed=2485233; DOI=10.1105/tpc.1.3.313;

RA Nielsen N.C., Dickinson C.D., Cho T.-J., Thanh V.H., Scallion B.J., Fischer R.L., Sims T.L., Drews G.N., Goldberg R.B.;

RT "Characterization of the glycinin gene family in soybean.";

RL Plant Cell 1:313-328(1989).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=cv. Dare; TISSUE=Leaf;

RX MEDLINE=89296499; PubMed=2740230;

RA Thanh V.H., Tumer N.E., Nielsen N.C.;

RT "The glycinin Gy2 gene from soybean.";

RL Nucleic Acids Res. 17:4387-4387(1989).

RN [3]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=cv. Shirotsurunoko;

RA Utsumi S., Kim C.S., Kohno M., Kito M.;

RT "Polymorphism and expression of cDNAs encoding glycinin subunits.";

RL Agric. Biol. Chem. 51:3267-3273(1987).

RN [4]

RP NUCLEOTIDE SEQUENCE.

RA Fukazawa C., Momma T., Higuchi W., Uidaka K.;

RT "Complete nucleotide sequence of the gene encoding a glycinin A2B1a subunit precursor of soybean.";

RL Nucleic Acids Res. 15:8117-8117(1987).

RN [5]

RP NUCLEOTIDE SEQUENCE.

RA Momma T., Negoro T., Uidaka K., Fukazawa C.;

RT "A complete cDNA coding for the sequence of glycinin A2B1a subunit precursor.";

RL FEBS Lett. 188:117-122(1985).

RN [6]

RP PROTEIN SEQUENCE OF 19-296 AND 301-480 (A2 AND B1a SUBUNITS).

RX MEDLINE=85030470; PubMed=6541652;

RA Staswick P.E., Hermodson M.A., Nielsen N.C.;

RT "The amino acid sequence of the A2B1a subunit of glycinin.";

RL J. Biol. Chem. 259:13424-13430(1984).

RN [7]

RP NUCLEOTIDE SEQUENCE OF 262-485.

RX MEDLINE=85030472; PubMed=6092376;

RA Marco Y.A., Thanh V.H., Tumer N.E., Scallion B.J., Nielsen N.C.;

RT "Cloning and structural analysis of DNA encoding an A2B1a subunit of glycinin.";

RL J. Biol. Chem. 259:13436-13441(1984).

RN [8]

RP NUCLEOTIDE SEQUENCE OF 1-36.

RX MEDLINE=90332420; PubMed=2377465;

RA Kitamura Y., Arahira M., Itoh Y., Fukazawa C.;

RT "The complete nucleotide sequence of soybean glycinin A2B1a gene spanning to another glycinin gene AlaB1b.";

RL Nucleic Acids Res. 18:4245-4245(1990).

RN [9]

RP PROTEIN SEQUENCE OF 115-136 AND 202-224 (A2 SUBUNIT).

RC STRAIN=cv. Kishinevskaaya-16;

RX MEDLINE=97054613; PubMed=8898910;

RA Shutov A.D., Kakhovskaya I.A., Bastrygina A.S., Bulmaga V.P., Horstmann C., Muntz K.;

RT "Limited proteolysis of beta-conglycinin and glycinin, the 7S and 11S storage globulins from soybean [Glycine max (L.) Merr.]. Structural and evolutionary implications.";

RL Eur. J. Biochem. 241:221-228(1996).

RN [10]

RP DISULFIDE BOND.

RX MEDLINE=85030471; PubMed=6541653;

RA Staswick P.E., Hermodson M.A., Nielsen N.C.;

RT "Identification of the cysteines which link the acidic and basic components of the glycinin subunits.";

RL J. Biol. Chem. 259:13431-13435(1984).

CC -!- FUNCTION: Glycinin is the major seed storage protein of soybean.

CC -!- SUBUNIT: Hexamer; each subunit is composed of an acidic and a basic chain derived from a single precursor and linked by a disulfide bond.

CC -!- SIMILARITY: Belongs to the 11S seed storage protein (globulins) family.

CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

CC -----

CC EMBL; X15122; CAA33216.1; -; Genomic_DNA.

DR EMBL; D00216; BAA00154.1; -; mRNA.

DR EMBL; Y00398; CAA68460.1; -; Genomic_DNA.

DR EMBL; X02806; CAA26575.1; -; mRNA.

DR EMBL; X02646; AAA33963.1; -; Genomic_DNA.

DR EMBL; X53404; CAA37480.1; -; Genomic_DNA.

DR PIR; A91341; FWSYG1.

DR HSSP; P04776; 1FXZ.

DR SMR; P04405; 26-479.

DR InterPro; IPR006045; Cupin.

DR InterPro; IPR006044; Seedstore_11s.

DR Pfam; PF00190; Cupin_1; 2.

DR PRINTS; PR00439; 11SGLOBULIN.

DR PROSITE; PS00305; 11S_SEED_STORAGE; 1.

KW Direct protein sequencing; Multigene family; Polymorphism; Seed storage protein; Signal; Storage protein.

KW SIGNAL 1 18

FT CHAIN 19 296 Glycinin A2 subunit.

FT PROPEP 297 300

FT CHAIN 301 480 Glycinin B1a subunit.

FT PROPEP 481 485

FT DISULFID 104 307 Interchain (between A2 and B1a chains).

FT VARIANT 103 103 G -> D.

FT VARIANT 318 318 N -> T.

FT VARIANT 331 331 I -> V.

FT VARIANT 413 413 K -> R.

FT CONFLICT 39 39 D -> G (in Ref. 5).

FT CONFLICT 39 39 D -> N (in Ref. 6).

FT CONFLICT 61 61 C -> S (in Ref. 6).

FT CONFLICT 117 117 R -> C (in Ref. 6).

FT CONFLICT 343 343 W -> S (in Ref. 6).

SQ SEQUENCE 485 AA; 54391 MW; 78BB459837F77AD8 CRC64;

Alignment Scores:

Pred. No.:	3,46e-96	Length:	485
Score:	1499.50	Matches:	297
Percent Similarity:	72.0%	Conservative:	71
Best Local Similarity:	58.1%	Mismatches:	94
Query Match:	54.0%	Indels:	49
DB:	1	Gaps:	6

US-10-728-323-3 (1-1524) x GLYG2_SOYBN (1-485)

QY 1 CGGAGCAGCGGAGGAGAGCGCTCCAGCTTCAGCGCTCAATGCCGAGACCTGAC 60

DB 20 ArgGluGlnAlaGlnGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProAsp 39

QY 61 AATCGCATTTGAATCAGAGCGGTTACATTGAGACTTGAACCCCAACCAACGAGGAGTTC 120

DB 40 AsnArgIleGluSerGluGlyGlyPheIleGluThrPAsnProAsnAsnLysProPhe 59

QY 121 GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCTCGCGCGCAACGCCCTTCGTAGGCCT 180


```
Db      60  GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 79
      181  TTCTACTCCAATGCTCCCGAGGATCTTCATCCAGCAAGGAGTACTTTGGTTG 240
      Db      80  SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyAsnGlyIlePheGlyMet 99
      QY      241  ATATTCCTCGTGTCTAGACACTATGAAGACCTTCACACAAAGGTGTCGATCTCAG 300
      Db      100  IlePheProGlyCysProSerThrTyrGlnGluProGlnGlnSerGlnGlnArgGlyArg 119
      QY      301  TCCCAAGACCAACCAAGACGCTCTCCAAGGAGAGACCAAAAGCCAACACAGCAACGAGATAGT 360
      Db      120  SerGlnArgPro-----GlnAspArg 126
      QY      361  CACCAGAAGGTGCACCGTTTCGATGAGGGTGATCTCATTCAGTTCCTCCACCGGTGTGCT 420
      Db      127  HisGlnLysValHisArgPheArgGluGlyAspLeuIleAlaValProThrGlyValAla 146
      QY      421  TTCTGCTCTACAAACGACACGACACTGATGTTGTTGCTGTTCTCTTACTGACACCAAC 480
      Db      147  TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerIleIleAspThrAsn 166
      QY      481  AACACGACAAACAGCTTGATCAGTTCCTCCAGGAGATTCAATTGCTGGGACACGGAG 540
      Db      167  SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
      QY      541  CAAGAGTTCTTAAGGTACCAACCAACAAAGACAGACAAAGCAGACAGAAAGCTTACCATAT 600
      Db      187  GlnGluPheLeuLysTyrGlnGln-----195
      QY      601  AGCCCATACAGCCCGCAAGTCAAGCTAGACAGAGAGCGTGAATTAGCCCTCAGAG 660
      Db      196  -----GlnGlnGlySerGlnSerGlnLysGly 205
      QY      661  CAGCACAGCCGACAGAGAACGAGCAGCAAGAGAGAAAGAAACCAAGCTGGAACATCTTC 720
      Db      206  LysGln-----GlnGluGluAsnGlnGlySerAsnIleLeu 218
      QY      721  AGCGGCTTACGCGGAGTTCCTGGAAACAGCTTCAGGTTGACGACAGACAGATAGTG 780
      Db      219  SerGlyPheAlaProGluPheLeuLysGluAlaPheGlyVal---AsnMetGlnIleVal 237
      QY      781  CAAACCTAAGAGCGCAGACCGAGAGTGAAGAGAGGAGCGCATCTGCAGTGAAGGGA 840
      Db      238  ArgAsnLeuGlnGlyGluAsnGluGluAspSerGlyAlaIleValThrValLysGly 257
      QY      841  GGCTCAGAACTTGAGCCAGATAGAAAGAGACGTGCCAGCAAGAGGAAATACGAT 900
      Db      258  GlyLeuArgValThrAlaProAlaMetArgLysProGlnGlnGluAspAspAsp 277
      QY      901  GAAGATGAA-----TATGAATACGATGAAGAGGATAGAAGCGGTGCAGGGGA 948
      Db      278  GluGluGlnProGlnCysValGluThrAspLysGlyCysGlnArgGlnSerLysArg 297
      QY      949  AGCAGAGGCGAGGGGATGTATTGAGACACATCTGCACCAAGTGCCTAAAAAGAAC 1008
      Db      298  SerArg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgGlnAsn 314
      QY      1009  ATTGGTATAGAAACAGATCCCTGCATCTCAACACCTCAAGCTGGTTCACCTCAAAAGTGC 1068
      Db      315  IleGlyGlnAsnSerProAspIleTyrAsnProGlnAlaGlySerIleThrAla 334
      QY      1069  AAGATCTCAACCTTCTAATCTTAGTGGCTGGACCTAGTGTGAATGAAATCTC 1128
      Db      335  ThrSerLeuAspPheProAlaLeuLeuLysLeuSerAlaGlnTyrGlySerLeu 354
      QY      1129  TACAGGAATCATGTTGTTCGCTCACTACAAACCAACGACACACAGCATCATATCGA 1188
      Db      355  ArgLysAsnAlaMetPheValProHisTyrThrLeuAsnAlaAsnSerIleIleTyrAla 374
      QY      1189  TTGAGGGGACGGGCTCACGTGCAAGTCGTGGACAGCAACCGCAACGAGTGTACGACGAG 1248
```

```
Db      375  LeuAsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGly 394
      QY      1249  GAGCTTCAAGAGGTCACTGCTTGTGTGGTCCACAGAACTTCGCGTCCCTCGAAGTCC 1308
      Db      395  GluLeuGlnGluGlyValLeuIleValProGlnAsnPheAlaValAlaAlaLysSer 414
      QY      1309  CAGAGCGAGAACTTCGAATACGTGCATTCAAGACAGACTCAAGGCCACGATAGCCAAC 1368
      Db      415  GlnSerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsn 434
      QY      1369  CTCGCGGTGAAACTCCGTCATAGATAACCTGCCGAGGAGGTGTTCCAAATTCATAT 1428
      Db      435  LeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnHisThrPhe 454
      QY      1429  GGCCTCCAAAGGAGCAGGACAGGAGCTTAAGACAAACCCCTTCAAGTTCCTCGTT 1488
      Db      455  AsnLeuLysSerGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuVal 474
      QY      1489  CCACCGCTCTCAGCAGTCTCCGAGGCGTGTGGCT 1521
      Db      475  ProProGlnGluSerGlnArgArgAlaValAla 485

RESULT 13
Q549Z4 SOYBN
ID Q549Z4_SOYBN PRELIMINARY; PRT; 485 AA.
AC Q549Z4_2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Proglycinin A2B1.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
OC Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Urade R., Nakatani H.;
RT "mRNA of Soybean Proglycinin A2B1 Subunit.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB113350; BAC78523.1; -; mRNA.
SQ SEQUENCE 485 AA; 54390 MW; 78BB459837F77AD8 CRC64;

Alignment Scores:
Pred. No.: 3,46e-96 Length: 485
Score: 1499.50 Matches: 297
Percent Similarity: 72.0% Conservative: 71
Best Local Similarity: 58.1% Mismatches: 94
Query Match: 54.0% Indels: 49
DB: 2

US-10-728-323-3 (1-1524) x Q549Z4_SOYBN (1-485)
QY 1 CGSGACAACCGGAGAGAACGCGTCAGTTCAGCGCTCAATCGCAGAGACCTGAC 60
Db 20 ArgGluGlnAlaGlnGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProAsp 39
QY 61 AATCGCATTAATCAGAGCGCGTTACATTGAGACTTGAAACCCCAACACAGGAGTTC 120
Db 40 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 59
QY 121 GAATGCGCGCGTCCCTCTCTCGTTAGTCTCCGCGCAACGCCCTTCGTAGGCCT 180
Db 60 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 79
QY 181 TTCTACTCCAATGCTCCCGAGGATCTTCATCCAGCAAGGAGTACTTTGGTTG 240
Db 80 SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyAsnGlyIlePheGlyMet 99
QY 241 ATATTCCTCGTGTCTAGACACTATGAAGACCTTCACACAAAGGTGTCGATCTCAG 300
Db 100 IlePheProGlyCysProSerThrTyrGlnGluProGlnGlnSerGlnGlnArgGlyArg 119
```



```
QY 301 TCCAAAGACCACCAAGACGCTCTCCAAAGGAGAGACCAAAAGCCAAACAGCAACGAGATAGT 360
Db 120 SerGlnArgPro-----GlnAspArg 126
QY 361 CACCAGAAGGTGCACCGCTTCGATGAGGGTGATCTCATTTGCCAGTTCCCAACCGGTGTGTGCT 420
Db 127 HisGlnLysValHisArgPheArgGluGlyAspLeuIleAlaValProThrGlyValAla 146
QY 421 TTCTGGCTCTACAAACGACACGACACTGATGTTGTTGCTGTTCTCTTACTGACACCAAC 480
Db 147 TrpTrpMetTyrAsnAsnGlnAspThrProValValAlaValSerIleIleAspThrAsn 166
QY 481 AACACCAACCAACGAGTTCATGATTCCTCCAGGAGATTCAATTTGGCTCGGACACGGAG 540
Db 167 SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
QY 541 CAAAGATTCTTAAGTTACGAGCAACAAACAGCAACAAAGCAGACGAGCAAGAGCTTACCATAT 600
Db 187 GlnGluPheLeuLysTyrGlnGlnGln----- 195
QY 601 AGCCCATACAGCCCGCAAGTCAGCCTAGACAAAGAGCGTGAAATTTAGCCCTCGAGGA 660
Db 196 -----GlnGlnGlySerGlnSerGlnLysGly 205
QY 661 CAGCACAGCCGACAGAACGACGAGCAAGAGAAAGAAACGAAGGTGGAAACATCTTC 720
Db 206 LysGln-----GlnGluGluGluAsnGluGlySerAsnIleLeu 218
QY 721 AGCGGCTTACGCGCGAGTCTCTGGAACAAAGCCTTCAGGTTGACGACAGACAGATAGT 780
Db 219 SerGlyPheAlaProGluPheLeuTyrGlnAlaPheGlyVal---AsnMetGlnIleVal 237
QY 781 CAAACCTTAAGAGCCGACAGCCGAGAGTGAAGAGAGGAGGACCATTTGTGACGTGAGGGGA 840
Db 238 ArgAsnLeuGlnGlyGluAsnGluGluGluAspSerGlyAlaIleValThrValLysGly 257
QY 841 GGCCTCAGATCTTGAGCCGACAGATAGAAAGACAGCTGCCGACGACGAGCAAGAAATACGAT 900
Db 258 GlyLeuArgValThrAlaProAlaMetArgLysProGlnGlnGluGluAspAspAsp 277
QY 901 GAAGATGAA-----TATGAATACGATGAAGAGGATAGAGCGCTGCAGGGGA 948
Db 278 GluGluGluGlnProGlnCysValGluThrAspLysGlyCysGlnArgGlnSerLysArg 297
QY 949 AGCAGAGCGAGGGGAATGTTTGAAGAGACGATCTGCACCGCAAGTCTAAAGAGAC 1008
Db 298 SerArg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgGlnAsn 314
QY 1009 ATTGGTAGAAACAGATCCCTGACATCTACACCCCTCAAGCTGGTTCACTCAAAAGTCC 1068
Db 315 IleGlyGlnAsnSerSerProAspIleTyrAsnProGlnAlaGlySerIleThrAla 334
QY 1069 AAGCATCTCAACCTTCTAATCTTAGTGCTTGGACCTAGTGTGCAATATGGAATCTC 1128
Db 335 ThrSerLeuAspPheProAlaLeuTrpLeuLeuLysLeuSerAlaGlnTyrGlySerLeu 354
QY 1129 TACAGGAATGCATTGTTGTCGCTACTACACCAACGACGACATCATATATCGA 1188
Db 355 ArgLysAsnAlaMetPheValProHisTyrThrLeuAsnAlaAsnSerIleIleTyrAla 374
QY 1189 TTGAGGGGACGGCTCACGTGCAAGTCTGGTGGACGCAACGCGCAACAGAGTGTACGACGAG 1248
Db 375 LeuAsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGly 394
QY 1249 GAGCTTCAAGAGGTACAGTCTGTTGTTGGTCCACAGAACTTCGCTCGCTCGGAAGTCC 1308
Db 395 GluLeuGlnGluGlyGlyValLeuIleValProGlnAsnPheAlaValAlaLysSer 414
QY 1309 CAGAGCGAGAACTTCCAAATACGTGGCATTCAAGACAGACTCAAGGCCACGATAGCCAAAC 1368
Db 415 GlnSerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsn 434
```

```
QY 1369 CTCGCGGTGAAACCTCCGTCATAGATAACCTGCCGAGGAGGTGGTTGCCAAATTCATAT 1428
Db 435 LeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnHisThrPhe 454
QY 1429 GCGCTCCAAAGGAGGACGAGGAGGAGCTTAAAGAACAAACCCCTTCAAGTTCCTCGTT 1488
Db 455 AsnLeuLysSerGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuVal 474
QY 1489 CCACGCTCTCAGCAGCTCTCCGAGGCGCTGTGGCT 1521
Db 475 ProProGlnGluSerGlnArgAlaValAla 485

RESULT 14
GLYGI_SOYBN
ID GLYGI_SOYBN STANDARD; PRT; 495 AA.
AC P04776;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DE 10-MAY-2005 (Rel. 47, Last annotation update)
DE Glycinin G1 precursor [Contains: Glycinin Ala subunit; Glycinin Bx subunit].
GN Name=GVL;
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
OC Glycine.
OX NCBI_TaxID=3847;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=cv. Dare; TISSUE=Leaf;
RX MEDLINE=92393391; PubMed=2485233; DOI=10.1105/tpc.1.3.313;
RA Nielsen N.C., Dickinson C.D., Cho T.-J., Thanh V.H., Scallion B.J.,
RA Fischer R.L., Sims T.L., Drews G.N., Goldberg R.B.;
RT "Characterization of the glycinin gene family in soybean.";
RL Plant Cell 1:313-328 (1989).
RN (2)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=cv. Dare; TISSUE=Leaf;
RX MEDLINE=89296498; PubMed=2740229;
RA Sims T.L., Goldberg R.B.;
RT "The glycinin Gyl gene from soybean.";
RL Nucleic Acids Res. 17:4386-4386 (1989).
RN (3)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=cv. Boninori;
RX MEDLINE=86041867; PubMed=2997720;
RA Negro T., Momma T., Fukazawa C.;
RT "A cDNA clone encoding a glycinin Ala subunit precursor of soybean.";
RL Nucleic Acids Res. 13:6719-6731 (1985).
RN (4)
RP NUCLEOTIDE SEQUENCE.
RA Utsuni S., Kohno M., Mori T., Kito M.;
RT "An alternate cDNA encoding glycinin Ala Bx subunit.";
RL J. Agric. Food Chem. 35:210-214 (1987).
RN (5)
RP NUCLEOTIDE SEQUENCE.
RA Urade R., Nakatani H., Nakano C.;
RT "mRNA of soybean proglycinin AlaB1b subunit.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Glycinin is the major seed storage protein of soybean.
CC -!- SUBUNIT: Hexamer; each subunit is composed of an acidic and a
CC basic chain derived from a single precursor and linked by a
CC disulfide bond.
CC -!- PTM: The precursor is post-translationally processed to form a
CC covalently linked Ala-Bx subunit complex.
CC -!- SIMILARITY: Belongs to the 11S seed storage protein (globulins)
CC family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
```

```
CC removed.
CC -----
DR EMBL; M36686; AAA33966.1; -; mRNA.
DR EMBL; X15121; CAA33215.1; -; Genomic_DNA.
DR EMBL; X02985; CAA26723.1; -; mRNA.
DR EMBL; AB113349; BAC78522.1; -; mRNA.
DR PIR; A23497; FWSYG2.
DR PIR; S10851; S10851.
DR PDB; 1FXZ; X-ray; A/B/C=20-495.
DR PDB; 1UCX; X-ray; A/B/C=20-495.
DR PDB; 1UD1; X-ray; A/B/C=20-495.
DR InterPro; IPR006045; Cupin_region.
DR InterPro; IPR007113; Cupin_region.
DR Pfam; PF00190; Cupin_1; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S_SEED_STORAGE; 1.
KW 3D-structure; Multigene family; Seed storage protein; Signal;
KW Storage protein.
FT SIGNAL 1 19 Glycinin Ala subunit.
FT CHAIN 20 306 Glycinin Bx subunit.
FT PROPEP 307 310 Glycinin Bx subunit.
FT CHAIN 311 490 Interchain (between Ala and Bx chains)
FT PROPEP 491 495 (By similarity).
FT DISULFID 107 317 D -> G (in Ref. 3).
FT CONFLICT 42 42 P -> S (in Ref. 3).
FT CONFLICT 108 108 F -> S (in Ref. 3).
FT CONFLICT 136 136 E -> G (in Ref. 3).
FT CONFLICT 360 360
FT TURN 30 31
FT STRAND 39 39
FT STRAND 43 47
FT TURN 48 49
FT STRAND 50 54
FT TURN 57 58
FT TURN 60 65
FT TURN 66 66
FT STRAND 68 74
FT TURN 76 77
FT STRAND 78 84
FT STRAND 89 95
FT STRAND 97 103
FT TURN 105 106
FT STRAND 134 137
FT TURN 138 139
FT STRAND 140 144
FT TURN 146 147
FT STRAND 149 154
FT STRAND 160 166
FT TURN 168 169
FT TURN 171 172
FT STRAND 181 183
FT TURN 191 192
FT HELIX 193 196
FT TURN 222 223
FT HELIX 226 233
FT TURN 234 234
FT HELIX 237 243
FT STRAND 255 257
FT TURN 317 318
FT STRAND 322 324
FT STRAND 333 335
FT TURN 336 338
FT STRAND 339 344
FT TURN 346 348
FT HELIX 350 352
FT TURN 353 356
FT STRAND 359 364
FT TURN 366 367
FT STRAND 369 375
FT TURN 376 377
FT STRAND 380 386
```

```
FT STRAND 388 394
FT TURN 396 397
FT STRAND 400 407
FT TURN 408 409
FT STRAND 411 414
FT TURN 416 417
FT STRAND 419 424
FT TURN 427 428
FT STRAND 429 435
FT STRAND 442 444
FT TURN 448 449
FT TURN 451 451
FT HELIX 452 454
FT HELIX 457 464
FT TURN 465 465
FT HELIX 468 476
FT STRAND 483 484
SQ SEQUENCE 495 AA; 55706 MW; 329CB0545B24D894 CRC64;

Alignment Scores:
Pred. No.: 6,1e-96 Length: 495
Score: 1496.00 Matches: 292
Percent Similarity: 70.9% Conservative: 73
Best Local Similarity: 56.7% Mismatches: 100
Query Match: 53.8% Indels: 50
DB: 1 Gaps: 6

US-10-728-323-3 (1-1524) x GLYGI_SOYBN (1-495)
QY 1 CGCGAGCAACCGAGAGAACGCGTCCAGCTTCCAGCGCTCAATGCGCAGAGACCTGAC 60
Db 23 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProAsp 42
QY 61 AATCGCATTGAATCAGAGGCGGTTACATTGAGACTTGAACCCCAACCAACGAGGATTC 120
Db 43 AsnArgIleGluSerGluGlyGlyLeuIleGluThrTrpAsnProAsnAsnLysProPhe 62
QY 121 GAATGCGCGCGTGCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCCT 180
Db 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 82
QY 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGAGGAGGGGATCTTTGGTTG 240
Db 83 SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyLysGlyIlePheGlyMet 102
QY 241 ATATTCCCTGGTTGCTTAGACACATATGAAGAGCCTCACACAAAGGTCGTGATCTCAG 300
Db 103 IleTyrProGlyCysProSerThrPheGluGluProGlnGlnProGlnGlnArgGlyGln 122
QY 301 TCCCAAGACCAACCAAGACGCTCTCCAGGAGAGAACCAAGCAACGACGAGATAGT 360
Db 123 SerSerArgPro-----GlnAspArg 129
QY 361 CACCAAGAGTGCACCGTTTCGATGAGGTGTATCTCATTCAGAGTCCCAACCGGTGTGCT 420
Db 130 HisGlnLysIleTyrAsnPheArgGluGlyAspLeuIleAlaValProThrGlyValAla 149
QY 421 TTCTGGCTCTACAACGACCAACGACACTGATGTTGTGTTCTTCTTCTACTGACACCAAC 480
Db 150 TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerIleIleAspThrAsn 169
QY 481 AACACGACACACGAGCTTGATCAGTTCCTCCAGAGATTCAATTTGGCTGGACACGGAG 540
Db 170 SerLeuGlnAsnGlnLeuAspGlnMetProArgPheTyrLeuAlaGlyAsnGlnGlu 189
QY 541 CAAGAGTTCTTAAGGTACCAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600
Db 190 GlnGluPheLeuLysTyrGlnGlnGluGln----- 199
QY 601 AGCCCCATACAGCCCGCAAAAGTCAGCCTAGACAAAGAGCGGTGAATTAGCCCTCAGGGA 660
Db 200 -----GlyGly 201
```


